THE UNIVERSITY OF ALBERTA

A STUDY OF THE EFFECT OF GROUPING FOR INSTRUCTION BY SEX ON SIXTH GRADE SCIENCE ACHIEVEMENT

by

© LESLIE LLEWELLYN THOMAS

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF MASTER OF EDUCATION

DEPARTMENT OF ELEMENTARY EDUCATION

EDMONTON, ALBERTA

SPRING, 1972

1972
144

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "A Study Of The Effect Of Grouping For Instruction By Sex On Sixth Grade Science Achievement", submitted by Leslie Llewellyn Thomas, in partial fulfilment of the requirements for the degree of Master of Education.

ABSTRACT

The purpose of the study was to examine the effect on science achievement of grouping students for instruction in science, using no other criterion than that of the sex of the student.

The population for the study consisted of eighty-seven grade six students located in four classrooms in two schools. The two classrooms in each school were instructed in science by the same teacher. Thus it was possible to designate one class for each teacher as experimental, and one class as control. The designation of the classes was left to chance.

Prior to the study the students were administered the STEP Form A science test, the Lorge-Thorndike intelligence test, and results were obtained for reading comprehension ability from the Ginn grade 6 reading test. These data, together with the students' ages were used as a group of covariates to control for student variability.

A special unit of instruction in science was introduced into the programme in science. This unit (Biogeography) was taught over an eight-week period. Two post-tests were administered at the end of the eight weeks. A STEP Form B science test was used as one post-test, and a multiple-choice test constructed by the investigator and the two cooperating teachers was used as the other.

A multiple linear regression analysis of covariance was used to analyse the data. The analysis was performed for the two criterion measures separately. A full group analysis and two sub group analyses

for the subjects of each teacher constituted the major analyses. The contribution of the covariates to the prediction of science scores was also tested, using a stepwise regression analysis.

The analyses of the hypotheses revealed little conclusive evidence to support the existence of significant differences in science achievement between the sub groups in the examination. One exception was that the students of teacher one, who were grouped for instruction by sex, were significantly superior in science achievement to those students of teacher one who were in mixed groups when the teacher post-test was used as the criterion.

Only previous science knowledge and intelligence were found to be significant predictors of both STEP post-test scores and the teacher constructed post-test scores. Reading did, however, contribute significantly to the prediction of STEP scores.

ACKNOWLEDGEMENTS

The writer wishes to express appreciation for all the assistance and cooperation he received in conducting and reporting the research project. In particular, I wish to express my gratitude to the following:

To N. M. Purvis, thesis committee chairman, for his guidance and encouragement, and to Dr. M. Horowitz, Dr. H. Kass, and K. G. Jacknicke, the other members of the committee, for their interest and advice.

To the Boards of the Peace River Public and Separate systems for their courtesy and cooperation in permitting my use of their classrooms, and to the pupils and teachers who took part in the study.

To Norm Watts and Roy Bailey, fellow students, for their continued encouragement.

Finally, to my wife Helen, and my family, for their patience and continued faith.

TABLE OF CONTENTS

CHAPTER PAGE

PAGE

LIST OF TABLES

TABLE		PAGE
XVI	SUMMARY OF RESULTS OF HYPOTHESES ON TWO CRITERION POST TESTS FOR TEACHER ONE	68
XVII	ANALYSIS OF COVARIANCE OF SCIENCE STEP TEST: TEACHER TWO HYPOTHESIS 1	70
XVIII	ANALYSIS OF COVARIANCE OF SCIENCE STEP TEST: TEACHER TWO HYPOTHESES 2 AND 3	71
XIX	SUMMARY OF RESULTS TEACHER TWO	72
XX	ANALYSIS OF COVARIANCE OF TEACHER POST TEST: TEACHER TWO HYPOTHESIS 1	73
XXI	ANALYSIS OF COVARIANCE OF TEACHER POST TEST: TEACHER TWO HYPOTHESES 2 AND 3	74
XXII	STEP-WISE REGRESSION ANALYSIS OF FOUR COVARIATES FOR STEP POST TEST	79
XXIII	STEP-WISE REGRESSION ANALYSIS OF FOUR COVARIATES FOR TEACHER POST TEST	81
XXIV	REGRESSION MODELS FOR TWO CRITERION MEASURES FULL GROUP ANALYSIS	103
XXV	REGRESSION MODELS FOR TWO CRITERION MEASURES TEACHER ONE	104
XXVI	REGRESSION MODELS FOR TWO CRITERION MEASURES TEACHER TWO	105

CHAPTER I

THE PROBLEM

INTRODUCTION TO THE STUDY

The elementary-school science programme for Alberta as outlined in the <u>Curriculum Guide</u> (1969) sets as one of the objectives of the programme that the students become active and dynamic investigators of science (p. 5). This objective is to be achieved by the development and use of process skills as they are defined in the guide.

The approach to the teaching of science in the elementary school requires a readjustment of traditional teacher-student roles. The guide makes the point

> . . . that for each suggested activity in this guide certain process skills are emphasized which can develop only through direct experience (p. 11).

Small group activities are suggested for situations involving experimental activities and it is implied that most concept development will take place in relation to concrete experiences.

It becomes necessary to arrange the science classes so that the pupils may work mostly in these small groups. The suggestion is made in the guide that the students be grouped randomly or by some form of ability grouping (p. 14).

BACKGROUND TO THE PROBLEM

Elementary schools generally organize classes for instruction

so that they contain both sexes and it is unusual that the sexes are separated for instruction, except in very special instances.

In grouping for science instruction involving small group activities, one could also consider the sex of the student as a criterion to be used in establishing the small groups.

Some writers have become increasingly concerned with the lack of attention being paid to the sex of the student when planning programmes of instruction. They feel, often for different reasons, that the manner in which instruction is made available to the individual student should take into consideration the possible different needs of the two sexes.

Peltier (1968) makes the point that elementary schools are sex-neutral institutions treating unequals as equals. Waetjen (1965), however, feels that elementary schools are feminine in emphasis, and as such that they handicap the development of boy pupils. Although they disagree as to cause, both writers feel that different approaches in instruction should be used for boys from those used for girls. They agree that the schools are prepared to accept conformity to the general standard (feminine standards in the elementary school), but that unorthodox ideas and novel attitudes are discouraged. As a result they see the initiative and creativity of boys being restricted by a system using instructional methods aimed at the development of conforming individuals. Lyles (1966) goes as far as to say that strong evidence exists to suggest that mathematics should be taught to boys in a masculine way and to girls in a feminine way.

Garai and Scheinfeld (1968) conclude from the results of a survey of studies of sex differences and related mental traits, that further research is necessary to determine how boys and girls differ with respect to skills, talents and cognitive styles. They recognize that the sexes share common abilities, but feel that the evidence indicates that there may be a feminine and masculine way of learning such subjects as mathematics and science. They suggest that part of the reason why girls in our society have shown an aversion to entering science fields in the past might be due to the fact that these subjects were generally taught by men, quite probably in a particularly masculine manner.

Epstein (1966), Farhner and Cronin (1967), and Parsley and Powell (1964) also consider pupil sex an important enough difference that it be given consideration when planning for instruction.

Yates (1966) concluded that differences related to the sex of the pupil, such as maturity, reading level and interests make investigation of grouping practices absolutely essential. He states quite explicitly

> There is still a good deal to be discovered about the extent to which it is possible or, on various grounds, desirable for boys and girls to pursue the same curriculum. Research is necessary, for example, to determine how far sex differences in attainment and interests in certain fields such as mathematics and science can be accounted for by cultural influences, and to what degree they are affected by differences in basic abilities and aptitudes (p. 135).

Scott (1964), studying science concept achievement of ten and eleven year olds, found little difference in terms of science concept -

development, but he did indicate trends which he felt required further investigation. He stated

> Since the interaction effect between age levels and sex categories appear to bear some relationship to science concept achievement, it may be implied that consideration be given to the way children are grouped for school activities (p. 16).

The question as to whether grouping by sex influences learning is open to continued discussion. This study was set up in an attempt to investigate whether science concept achievement at the grade six level is facilitated when the pupils are arranged for science group work according to sex.

Investigations have been carried out to determine whether segregation by sex in whole schools has any effect on learning.

Durante (1967), Farhner and Cronin (1963), Herman and Criscuolo (1968), and Wisethal (1965) all report studies of grouping by sex for instruction in whole school situations. The consensus among them is that differences in results do occur as a result of grouping by sex. While these results indicate possible advantages of this kind of grouping, they also feel that more investigation is indicated.

Lyles (1966), Stanchfield (1965) and Hurd (1934) also report grouping experiments in segregated and non-segregated classes. In general the results show increased achievement for boys in segregated classes. Hurd and Lyles studied science achievement in the high school and elementary school respectively.

This study restricts the separation of the students by sex to those situations within science classrooms which require involvement

in group work.

HYPOTHESES

Knowledge of how students are grouped for instruction does not contribute significantly to the ability to predict science achievement scores

1 between all experimental subjects and all control subjects

1a between experimental females and control females

1b between experimental males and control males

2 between experimental females and experimental males

3 between control females and control males

when science pre-test scores, reading comprehension ability, intelligence and age are controlled.

DEFINITION OF TERMS

For the purposes of this study the following definitions will be used:

Experimental subjects

Students organized for science group work, within classes containing both sexes, so that the groups are all-boy and all-girl groups.

Control subjects

Students organized for science group work within classes so that each group includes both boys and girls.

Intelligence quotient

The subjects' score on the Lorge-Thorndike Intelligence Test Form 1, Level D.

Reading ability

The score obtained from Part Three (Comprehension) of the Sixth Reader Achievement Test of the Ginn Basic Reading series.

Pre-test scores

Converted scores obtained on Level Four, Form A, of the Sequential Test of Educational Progress Science Test.

Science concept achievement

Converted scores obtained on Level Four, Form B, of the Sequential Test of Educational Progress Science Test.

Scores obtained on the teacher constructed post-test, based on the six Big Ideas of the Biogeography teaching unit.

LIMITATIONS

In interpreting the data of this study the following limitations should be borne in mind:

1. Science achievement cannot be considered to be only that which can be measured using multiple-choice tests. Subjective evaluation of subjects' responses in day-to-day reporting proved beyond the scope of this study, restricting it in this respect.

2. The instrument used to obtain part of the post-test data was constructed by the investigator and the cooperating teachers. The validity of the instrument is unknown, but efforts were made to ensure

that the test measured those concepts associated with the Biogeography unit.

3. The sample for the study was selected from a population of grade six students in Peace River. It was assumed that this sample was representative of urban children at this level.

4. While the use of male teachers in the study does reduce variability, it might also introduce into the data some factor as a result of the interaction between the men teachers and the boy subjects.

SIGNIFICANCE OF THE STUDY

Increased concern with individualizing instruction, and continued controversy surrounding learning and its relationship to sex, suggest the need to gather information regarding the various organizations possible within the classroom. Certainly there is evidence that girls on this continent have not entered science careers as readily as boys. A survey of registered scientists in the USA for 1964 revealed that of 223,854 registered scientists 92.3% were men.

While it is certainly true that many factors influence such figures, including the cultural expectations for girls, it is possible that science programmes in our schools do not cater to girls in the kinds of activities they provide.

Ideally, students of science in the elementary school should be deeply involved in those activities which the Curriculum Guide calls the 'process' of science. How the teacher caters to this need for involvement is of major concern. As Dunfee (1967) says

> No science programme, how ever well conceived and organized, no matter how firmly based on the opinions of experts and knowledge about children can be more successful than its implementation in the classroom . . . How his group is organized for instruction, how he teaches science . . . becomes the focus for concern (p. 28).

Trends indicated by this study might suggest further studies of a similar nature, perhaps over a longer time period. Finally, some indication of the value of grouping by sex at other grade levels might be investigated if these results show this to be desirable.

THE EXPERIMENTAL SETTING

The population from which the sample was drawn consisted of the six grade six classrooms of the Public and Separate school systems in Peace River town. The sample consisted of eighty-seven of these students in four of the classrooms. The sample was divided into experimental and control groups randomly. The two classes in the experimental group contained student groups which were all-boy or all-girl groups. In the control classes the students were in mixed sex groups.

A special unit of instruction was introduced into the classrooms for an eight week period. Eleven activities were selected from the unit, and presented to the students over the period of the study.

A pre-test was administered to measure previous science knowledge prior to the implementation of the study. Data concerning IQ, reading ability and age were also collected.

A post-test consisting of two criterion tests was administered at the end of the eight week period. The two tests were the STEP and the test designed for the study (see page 30).

The design of the study was affected by the fact that the researcher was closely involved professionally with the teachers in the study prior to setting it up. The expectation of continuing close professional involvement also placed some restrictions on the design.

OUTLINE OF THE REPORT

A review of the relevant literature and research will be presented in Chapter II. Chapter III describes the experimental design of the study, including descriptions of tests and of procedures.

The results of the data analyses are presented in Chapter IV. The final chapter, Chapter V, includes a summary and a discussion of the findings, and the implications for further educational research.

CHAPTER II

REVIEW OF RELATED LITERATURE AND RESEARCH

The purpose of this study is an attempt to show that science achievement, at the grade six level, would be improved for both boys and girls if instruction in science was developed through small group activities presented to all-boy and all-girl groups. The implication is that interaction between boys and girls in the same group tends to inhibit the science achievement of both sexes.

Many factors are included in the total picture of academic disparity between the sexes which some researchers claim exists in the elementary school. This review, as far as possible, limits itself to those studies which discuss factors having some direct bearing on elementary school science achievement.

BACKGROUND AND INTEREST IN SCIENCE

That the difference in background of boys and girls in science is not often argued against is evident from an examination of the literature. A number of studies show that significant differences in background and interest in science between the sexes can be measured.

Swan (1967) examined the science achievement of more than three thousand sixth grade students. He found that students scoring higher on a science interest survey also scored higher on a science achievement test. On both the interest and achievement tests the

results favoured the boys significantly. Uhlhorn (1963) investigated the science backgrounds of fourth, fifth and sixth grade students by using a science experience inventory he developed. He found that the boys had a significantly greater number of science experiences than did the girls.

Smock and Holt (1962), reporting a study of grade one students' curiosity motivation, show males to be significantly more curious than females at this level. They found that females tend to maintain rigidity of action when placed in situations involving choices among different actions. This rigidity might have implications in the experimental situations which are a part of elementary science.

Lewis (1967) administered fact content tests in science to three hundred high school boys and girls. He reported differences between the sexes at the 5% level of significance, showing boys to be more interested in science than girls.

Perrodin (1966) studied the attitudes of grade six students to school subjects. His findings showed boys considered science a more important subject than did the girls.

Rowland (1968) used a background experience scale with one hundred forty-four boys and one hundred forty-four girls in grade six. He reported significant differences at the 1% level favouring boys in terms of their science backgrounds.

Garai and Scheinfeld (1968) report a comprehensive survey of four hundred seventy-four studies and articles which examine sex differences and related mental and behavioural traits. They conclude, as

does Rowland, that very different interests and levels of achievement exist between the sexes. In particular they feel that science instruction in the elementary school does not cater to the interests or the needs of girls. They point out that

> . . . further intensive research is required to identify more adequately the specific different as well as similar skills, talents and cognitive styles which boys and girls utilize in the learning of various science subjects (p. 277).

Garai and Scheinfeld go on to say that boys might tend to attain higher levels of achievement in all-male classes, particularly in those areas such as science, in which they surpass girls. And further, they point out that science teaching must be made stimulating and interesting for both sexes if optimum results in science achievement are to be attained.

The differences in interest in science between the sexes manifest themselves relatively early in the child's school life. Cooley (1964) found that boys and girls interested in science had different basic patterns of interests. Investigating the vocational interests of children, he found that boys wanting science careers crystallized their interest somewhere between the age of ten and fourteen.

It would appear that boys have greater opportunity for entering science careers than girls. However, increasing female involvement in the fields of science in the USSR and other countries of Europe indicate that much of the reason for women avoiding science careers might be in the process of being overcome.

The general consensus of opinion among writers and researchers

is that boys exhibit greater interest in things scientific than do girls.

SCIENCE ACHIEVEMENT

Research into differences in science achievement between the sexes, although not without disagreements in the findings, generally leans towards a conclusion that, on the average, boys achieve more than girls.

DeLorenzo (1963) studied grade six science achievement when the teaching was done by regular classroom and 'special' teachers of science. His results show boys to be superior in mean science achievement on both his pre and post-test results. He reports differences in achievement significant at the 0.05 level.

Barker (1965), studying the effect of a 'discovery table' on the science achievement scores of ten and eleven year olds, found that the girls were consistently scoring lower than the boys. This study was of particular interest in that the subjects had been exposed to no science instruction at all prior to the study.

In a study of the science achievement of gifted boys and girls at the grade six level, Becker (1964) found that the gifted boys were significantly superior to the gifted girls ($p > 0.01$). In this study the students were grouped according to ability. Becker reports that the boys in the non-gifted group also scored higher than the girls in that group.

Atkinson (1931), in an early study of general science achieve-

ment at the junior high level, tested one hundred nineteen cases. He found slight differences favouring boys in science achievement, even though the girls scored higher on the intelligence tests.

Koutnic (1964) designed a battery and bulb test to be administered to students at the grade six level. His test was an on/off problem using simple circuits. His findings showed boys out-performing girls throughout the study of eighty-seven subjects. However, Koutnic felt that the difference in ability was more related to background differences than to differences in intellectual functioning.

The science achievement of boys and girls differs at all grade levels. It is possible that the achievement of girls in science is related to their greater interest in the biological sciences than in the physical sciences.

INTELLECTUAL DIFFERENCES

Whether the difference which exists in terms of science achievement between the sexes is the result of basic intellectual differences is still open to conjecture. Although much evidence exists showing no differences between the sexes in general intelligence, the point has been made (Jones, 1969, p. 6) that most intelligence test makers discard those items which tend to discriminate between the sexes, thus designing out of the tests any ability to relate performance to sex.

Bennett, Seashore, and Wesman (1956) have reviewed the Differential Aptitude Test. They concluded that the norms for the sexes

differed. The boys were higher in terms of spatial relations and mechanical reasoning, while the girls outscored the boys on language usage, spelling and clerical skills.

Wisenthal (1965) has studied differences between the sexes in attitudes and attainment of elementary school children. In his report he draws attention to the study carried out by the Scottish Council on Education in 1949. All eleven year olds enrolled in the schools in Scotland were administered the Terman-Merrill revision of the Stanford-Binet IQ test. The overall findings indicated a significant difference in scores between the sexes, with the boys scoring higher.

Weschler (1958), who developed the WISC and WAIS tests of intelligence, contends that while the sexes might score higher than each other on any given test, they have, in fact, different types of intelligence. He says

> . . . women seemingly call upon different resources or different degrees of like abilities in exercising whatever it is we call intelligence . . . but our findings confirm . . . that men not only behave, but 'think' differently from women (p. 148).

However, if one accepts that little difference exists between the sexes in general intelligence, one needs to consider that it is possible that the sexes do, as Weschler says, 'think' differently in such fields as mathematics and science.

Waetjen, in reviewing recent research on science teaching (1965), reports some evidence indicating that boys have different analytic thinking skills than do girls in terms of science. Peltier (1968) supports this contention in a discussion of sex differences

in the elementary school.

Sigel (1965) also considers that males and females perform on different levels on cognitive tasks. Philip (1949) decided that boys and girls use different abilities for the solution of mechanical problems.

The ability of elementary school children to use mental models in the explanation of natural phenomena was examined by Anderson (1965). He interviewed one hundred eighty subjects, demonstrating different phenomena to them. He found that boys were more likely to offer atomistic models in explanation, but he noted that the difference between the sexes was such as to provide limited evidence.

Within this whole area of intellectual activity, the ability to transfer knowledge to new situations appears to be directly related to the ability to achieve in school. Studies which investigate transfer abilities have direct implications for science learning in the elementary school, since much of the concept development in science takes place in situations involving the manipulation of information and its transfer to situations of an investigative nature.

Lyles (1966) found differences in ability to transfer knowledge between five hundred subjects in control and sex-segregated classes. His investigation covered grades one to six in the area of the subjects' science achievement. He found boys able to transfer knowledge and apply learned skills in science more readily than girls.

Kostic (1954) also noted transfer ability differences related to the subjects' sex when investigating differences in the learning

process. Although his study involved high school students, a point of interest was that boys apparently transfer information more readily than girls even when the medium of instruction is entirely unfamiliar. In this case the boys were taught home economics.

Hill and Dusek (1969) examined the achievement expectations of grade three and four students. They studied forty-eight boys and forty-eight girls to determine if differences related to the sex of the student existed in terms of their achievement expectations. They noted differences at all levels of their investigation which led them to conclude that

> There is an increasing body of evidence that sex differences in achievement behaviour and intellectual functioning may be greater than has been realized (p. 551).

MANIPULATIVE SKILLS

The ability to deal with the material aspects of scientific investigations has a bearing on the degree of involvement individuals reach in this respect. While it has been recognized that girls generally have finer coordination skills in terms of cutting out and colouring in the early elementary grades, there is evidence to show that boys are more prepared to manipulate objects when examining them than are girls.

Goldschmid (1967), in studying various conservation skills, found that boys in general in grade one and two appear to be better conservers than girls. He found significant differences between boys and girls in their ability to conserve substance and discontinuous

quantity, and also that the boys' total scores were superior to the girls'. He suggested that these differences might be the result of the boys having more opportunity to manipulate concrete objects during their play activities. He felt that the boys also examined transformations related to concrete objects more than girls.

Fogelman (1970) also examined Piagetian tests in terms of sex differences. His subjects were six and seven year olds. He determined that the boys performed the tasks better when they were presented to them in manipulative situations. The girls, on the other hand, seemed to do better when they were permitted to remain passive and watch. He concluded that the Piagetian tests failed to discriminate accurately between conservers and non-conservers since the mode of presentation gave an advantage either to the girls or to the boys, depending on which mode, passive or manipulative, was used.

Walberg (1967) has studied the science interests of boys and girls in high school. He stated that in high school physics, boys would be more attracted to a method of experimentation, while girls would be more interested in a method of instruction employing discussion and questioning. He felt that the boys learned more from the manipulation of physical objects in the experiments than did the girls.

Wisenthal (1965) has examined the achievement of boys in all-boy schools and girls in all-girl schools. Among various findings he reported that girls tend to show up better on tests involving verbal situations than boys do. However, it was noted that the boys did consistently better on tests involving performance items.

While the evidence seems to indicate boys' willingness to become involved with the physical manipulation aspects of scientific investigations, there is little doubt that the girls become the reporters for the group, either keeping notes and records, or preparing the final reports and presenting them. Differences in skills in communication, thus have important implications for small group work in science.

COMMUNICATION SKILLS

The ability to communicate the results of investigations is an important part of the process of science. There is evidence which indicates that girls in the elementary school probably have superior skills in the field of reporting the results of investigations. This evidence has implications for the teacher of science in the elementary school.

Garai and Scheinfeld (1968) show that many of the studies they review indicate that girls have superior language and verbal skills. Clark and Willis (1959) found differences significant at the 1% level between boys and girls in language skills. Studying almost seventy thousand subjects in grades three, five and eight, they reported results showing girls to be significantly superior in language skills measured with standardized tests. Chronological and mental age were held constant in this study.

DeRoche (1966) investigated the effect of creative exercises on creative thinking in terms of mastery of a unit of science at the

grade six level. His findings showed boys in both the experimental and the two control groups to have superior ability in science to the girls in the groups. However, his findings also showed that the girls exhibited greater fluency in science than the boys. All results were claimed to show significant differences.

An interesting study of masculinity and femininity as a personality variable, and its relationship to student achievement in the elementary school was carried out by Lamkin (1967). He examined intersex and intrasex differences in terms of fourteen variables. While his results showed greater intrasex differences than intersex differences, they also showed that the girls as a group were significantly superior to the boys in language achievement (0.001 level).

An extended study of the difference in learning patterns of boys and girls in grade one led Stanchfield (1968) to conclude that girls at this level had superior verbal ability. While it is true that these patterns change, there is little doubt that the girls still out-score the boys in verbal ability at the grade six level. Stanchfield's study, lasting six years, concentrated only on reading ability. However, it was noted by all the teachers involved that gross differences in ability to verbalize thoughts existed between the boys and the girls. These differences became apparent as soon as the boys were grouped together for instruction. Stanchfield felt that the differences in communication skills which exist between the sexes are hidden in mixed sex classes.

However, not all of the researchers investigating differences

in language skills agree as to the reasons why these differences exist. Terman and Tyler (1954) consider that girls tend to have superior teacher ratings in general achievement, but that they do not always show this superiority on tests. However, they do recognize that the contention that girls show superiority in language skills is supported by both teacher ratings and standardized tests.

Maccoby (1968) also agrees that girls show superior language skills to boys, but the feeling is expressed that the differences are not as marked as are sometimes claimed. Cardon (1968) takes a more extreme position than this, stating that the academic disparity which exists between boys and girls in the elementary school might even be the result of some kind of genetic language superiority possessed by the girls.

One can conclude from the evidence that there are significant differences in the elementary school favouring boys in terms of science achievement, science interest and background, and that differences favouring girls with respect to communication skills are also indicated. Some evidence also exists favouring the boys with more manipulative ability, and cognitive processes which tend to improve their performance in science situations.

With these differences in existence, grouping by sex for instruction in science has features making it attractive to the teacher of science in the elementary school. Some studies of grouping practices have been carried out, although much of the data cited as showing differences between the sexes is usually collected as a result

of studies of other aspects of learning. Grouping studies usually restrict themselves to all-class or all-school sex groups.

GROUPING

Scott (1964) suggested that consideration be given to the way children are grouped for science instruction. Scott investigated science achievement and cognitive functioning with ten and eleven year old students. He reports some interaction between age and sex categories and science achievement, and says

> It is suggested that group activities by children can produce good results when students in these groups are compatible in various ways (p. 16).

Margenstren (1966) reports a number of articles on grouping in the elementary school. While grouping by sex is not discussed at length, Goldberg and Passow point out that

> Various kinds of grouping . . . can be used effectively when they are designed to implement planned variation in method and content.

This implies that the organization of classes into small groups consisting of one sex would be a possible alternative to other grouping methods for the science teacher.

Farhner and Cronin (1963) report the results of a pilot project carried out in grade eight. They feel that individual differences might be more important than sex differences in terms of the learning ability of grade eight students. However, they do point out that grouping by sex is

> . . . just one of the many proposals for intelligent scheduling of boys and girls

and that teachers

> . . . should recognize bv now that sex differences in learning are numerous and genuine (p. 17).

Herman and Criscuolo (1968) organized four grade one classes so that two classes were mixed by sex while the other classes were all-boy and all-girl classes. Their findings indicated that grade one boys had special needs which needed catering to.

Wisenthal (1965) studied sex differences in attitudes and attainment in English junior schools. These schools were sex-segregated and compare in student age to elementary schools in this province. His findings did not show that boys benefitted from isolation in all-boy schools. On the other hand, Wisenthal's results supported the contention that girls do benefit from such isolation. He concluded that basic psychological differences do exist between boys and girls in junior schools.

Hillerich (1967) investigated interclass grouping in the elementary school. He reviewed research in this field and came to the conclusion that children cannot be 'homogenized' by ability or by achievement, but that one possibility being somewhat ignored is that of grouping by sex. He says

> . . . perhaps some future efforts at innovative grouping--and this is not really new either--might be devoted to grouping in the one way that kids can be alike, i.e., by sex (p. 8).

SUMMARY

The communication of the results of inquiry are essential to the process of science. Evidence indicating differences in language skills favouring girls would influence this communication in mixed-sex groups. It is possible that the girls in the groups would tend to become the reporters for those groups. The boys' involvement in reporting activities might be inhibited by a lack of communication skill necessary for these activities, and thus not provide them with the opportunity to develop reporting skills. In single sex groups the students would be required to give more attention to the total spectrum of science skills.

On the other hand, boys apparently show some superiority in their ability to transfer information in problem situations. They also have some apparent 'tinkering' skills (Walberg 1967, p. 115). If the girls do indeed have less well developed skills in these areas, they might in turn be inhibited in their involvement in the small group experimental work which is so much a part of the daily work in science in the elementary school.

Grouping by sex for science instruction would allow the students to develop all of the process skills which the Curriculum Guide lists as being essential in learning science, and in effect allow the teacher of science to plan for developmental work in those areas which the differing needs of the boys and girls make necessary.

Since the science programme is such that much group work is necessary it would appear that an opportunity to utilize grouping by

sex within the classroom exists. The needs of the two sexes could be met in science with the least departure from the co-educational nature of our elementary schools.

CHAPTER III

DESIGN OF THE STUDY

SAMPLE

The population consisted of the grade six enrollments in the Peace River Public and Separate school systems. This comprised a total of about 150 children in six classrooms located in three schools. Each school operated two grade six classrooms. The classrooms were made available for the study following requests to the two school Boards.

The sample subjects were those students enrolled in four of the classrooms, two rooms in each of two schools. The two classes not included in the study had been segregated into an all-boy class and an all-girl class during the whole of the school year. Since the purpose of the study was to examine the effect of grouping by sex within classes containing both sexes, these two classes were excluded from the study as not meeting the criterion of having both sexes in one room.

The total number of subjects involved in the study was eighty-seven.

The science instruction for the four classes in the study was carried out by two men teachers. Since each teacher was responsible for two classes in the same school, it was felt that the choice of experimental and control classes could be made by the toss of a coin.

This was done, and as a result each teacher was required to present the same science experiences to both a control and an experimental class.

TABLE I provides a classification of students by sex and by treatment.

TABLE I

CLASSIFICATION OF PUPILS BY SEX
AND BY TREATMENT

Sex	Experimental	Control
Boys	20	22
Girls	21	24
Totals	41	46

INSTRUMENTATION

Science Pre-test

The Sequential Test of Educational Progress, Science Form 4A was chosen for the pre-test. A search was made for an instrument which would measure those process skills considered an integral part of science learning. Of available materials, this test seemed most suitable for the purposes of this study. Inquiry skills in six areas are measured by the test:

1. Ability to identify and define scientific problems.
2. Ability to suggest or screen hypotheses.
3. Ability to suggest valid procedures.
4. Ability to interpret data and draw conclusions.
5. Ability to reason quantitatively and symbolically.
6. Ability to evaluate critically claims or statements made by others.

The test content places emphasis on Biology (40%), and considering the nature of the instructional material used in this study, was felt to be a suitable instrument in this respect.

The authors of the test report that reliability estimates were calculated for the middle grade group on Form 4A of the test. They report a reliability coefficient of 0.80 at the Grade 5 level.

Some examples of questions from the tests are given in Appendix A.

Intelligence Quotient

The Lorge-Thorndike Intelligence Test, Level D, is administered to all grade six pupils in the Peace River systems in April of each year. The results for the 1970 testing were used to control for student IQ variability.

Reading Ability

The students in the study use the Ginn reading programme, with tests administered to them at the end of each grade level. Since the science tests used in the study were heavily oriented toward reading skill, it was felt that comprehension ability would play a large part in student performance on the tests. The comprehension total from the Ginn test was used to control for variability in the students' comprehension ability. The reading test results for June, 1970, were used.

The age of the students was also considered in the analysis of the data. TABLE II shows unadjusted means for the control and experimental groups for pre-test science scores, reading scores, intelligence, and age.

TABLE II

UNADJUSTED MEANS PRE-TEST COVARIATE SCORES

	Experimental		Control	
	Boys	Girls	Boys	Girls
Science	264.4	264.3	271.1	266.4
Reading	78.2	79.3	85.1	84.8
Intelligence	102.3	103.2	107.9	101.4
Age (in months)	144.4	143.1	143.2	143.5

Science Post-tests

Since the STEP series of tests has matched forms, it was possible to use Form 4B as a post-test. Comparison of results between the forms required that raw scores were converted (Manual and Key for scoring Science Test Form 4).

The investigator and cooperating teachers felt that some valuable information could be gained by testing the hypothesis, that knowledge of how students are grouped for science instruction does not contribute significantly to the ability to predict science scores on two criterion measures. It was thought that the use of a teacher-designed test might yield results having more relevance to the classroom situation than the STEP results would have. Therefore test questions were designed based on the materials used in instruction. The test consisted of thirty multiple-choice items designed to test for

understandings derived from the Biogeography unit.

The questions were developed by testing the grade six classes not included in the study, but who had been receiving instruction with the same activities from the Biogeography unit. The questions were constructed, used to test the students, and then modified as was felt necessary. It was not possible to analyse the results mathematically at this stage, but an item analysis was performed on the results of the final test with the subjects in the study.

The Kuder-Richardson Formula 20 was used to estimate the reliability. The results indicated a low internal consistency of 0.52. As Ferguson points out, a test will tend to have high reliability if the items on the test have high correlations with each other and are measures of much the same attribute (p. 380).

The average biserial correlation of 0.67 indicates a low degree of correlation which might be the result of the test items measuring differing underlying attributes, perhaps such as the ability to hypothesize, make inferences, and recall information. The short length of the test would also affect the reliability estimate since longer tests have higher reliability coefficients than similar tests with fewer items (Ferguson, page 381). The average item difficulty was 0.61 which approximates the ideal of an average difficulty index of 0.50 for multiple choice items.

Both parts of the testing programme were administered to the subjects by the classroom teachers in order to maintain as natural a classroom atmosphere as possible. These tests were administered during

final week of the school year. The STEP was administered first in all four classrooms. The grading of the tests was carried out by the investigator.

INSTRUCTIONAL MATERIALS

Since it was essential that all the subjects in the study be exposed to the same science experiences, and since it was desirable to have as much of the science work consist of small group activities as possible, it was decided to use a special unit of work for the duration of the study.

An examination of available materials led to the choice of the unit BIOGEOGRAPHY from the Science Research Associates programme Learnings in Science (D. Parker, D. Stotler, and A. Rand). Activities were chosen from the three levels of each of the six Big Ideas in the programme. After discussion with the cooperating teachers, a total of eleven activities were chosen to be used during the eight weeks of the study.

While it was recognized that if time became a factor each pair of classes might not cover exactly the same activities, it was considered unnecessary to correct for this possibility. As long as each of the teachers made sure that his control and experimental classes covered the same experiences, little overall variability would be introduced into the study in terms of the content covered.

The criteria for choosing activities included appropriateness for small group activity, variety of reporting methods used, use of

both living and non-living materials, and ease of obtaining the materials required. An attempt was made to include activities which required the students to present their results in different ways. In some instances the results were presented in the form of tables or graphs, in other instances written reports were required.

Each activity was set up on a card which outlined the problem under study. The materials required were listed on the card, and a general description of how to set up the investigation was also given. Sheets were distributed to the groups for each investigation.

Appendix C lists those activities from the unit which were used in the study. One example activity is also given.

To facilitate the study, and to control for any variations which might occur as a result of the experimental nature of the material, the investigator supplied instructional materials as required to the four classrooms in the study.

PROCEDURE

Since the purpose of the study is to investigate the effect of grouping by sex on science achievement, the treatment consists of establishing different sex groupings in the experimental classes and the control classes.

The students were organized for instruction into groups of three or four. This size group seemed to be optimum in the circumstances for the kind of instruction made necessary by the unit used. The Curriculum Guide (1969) suggests groups of two or three (p.14).

However, it was felt that groups of three or four would mean less difficulty in maintaining the supply of materials and equipment.

The criterion used for grouping in the control classes was that each group should contain at least one girl and one boy. It was recognized that student absence during the period of the study might require that regrouping be undertaken in the control classes to maintain the mixed-sex grouping criterion. As no system could be devised to preplan for this eventuality, it was decided by the investigator that the classroom teacher should deal with the problem if and when it arose.

In the experimental classes the students were organized into groups of three or four for instruction. The criterion used in establishing these groups was that each group must consist of all boys or all girls. Student absence would not require that student regrouping take place during the study.

The division of the classes into instructional groups was undertaken by the classroom teachers according to the criteria outlined previously. The choice of students to be placed in particular groups was left up to the teacher since this is normally the way groups are established in the classroom.

The teachers cooperating in the study were both men, and although it was recognized that this fact placed certain limits on the study, it was also felt that the variable of teacher sex would be controlled, and differences in interaction between students and teachers would be reduced.

Since each teacher instructed both an experimental and a control class, variations in teacher performance would be expected to influence the results in both classes, thus controlling for the differences in teacher performance which might result from using more than one teacher in the study.

ANALYSIS

The hypotheses were examined by means of the linear regression model for one-way analysis of covariance. Since the class groups used in the study were existing classes, allocation to the experimental and the control groups was made on no criterion other than class membership. Ferguson (p. 326) points out that covariates not controlled experimentally may influence the ability to predict criterion scores. However, the analysis of covariance technique does permit these concomitant variables to be controlled statistically.

Kelly (p. 225) suggests that the use of the multiple linear regression procedure allows the investigator to ask whether the additional knowledge of the treatment of the group significantly improves the prediction of the criterion variable over and above that which the concomitant variables predict.

Kelly points out that more than one concomitant variable can be dealt with by using the linear regression model. The statistical control of the covariates is achieved by the addition of a vector for each covariate to be controlled to the full and restricted models.

Intelligence, age, previous science knowledge, and reading

comprehension ability are relevant concomitant variables, which it was felt, influence the ability to predict science achievement. These variables are treated as a group and are controlled for statistically in the analyses.

The linear regression approach examines the ability of two linear regression models to predict the criterion, namely, science achievement. The first, or 'full', model contains group membership vectors of the subjects by treatment as well as the vectors of the covariates.

Full model:

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + a_9X_9 + a_{10}X_{10} + e$$

Where:

X_3 is the criterion to be predicted

X_2, X_5, X_7, X_8 are the concomitant variables of previous science knowledge, intelligence, age, and reading comprehension ability respectively.

X_9 and X_{10} are the vectors for the treatment groups.

a equals the unknown weights which make the linear statement true (Kelly, p. 56).

e equals the difference between the observed score and the predicted score on the criterion measure (Kelly, p. 120).

The second or 'restricted' model contains only the unit vector and the covariates

Restricted model:

$$X_3 = A_o U + a_2 X_2 + a_5 X_5 + a_7 X_7 + a_8 X_8 + e$$

Where the symbols are defined as above, but the weights a, and the error e, are different values in the two equations.

Any statistically significant difference between the ability of these models to predict the criterion score is regarded as being due to the difference in group membership (treatment). These differences would be presumed to be present over and above the effects of the covariates, since the covariates appear in both models.

The analysis of the relationship between predictor variables and predicted score produces a 'line of best fit.' That is, a line is drawn statistically which will minimize the variance between observed and predicted scores on the criterion. The procedure determines the smallest discrepancies between the sums of squares predicted and the sums of squares observed (Kelly, p. 125).

The squared multiple correlation coefficients (R^2) represents the proportion of criterion variance accounted for by the set of predictor variables. The significance of the difference between the R^2 for the full model and the R^2 for the restricted model, and thus the ability to predict science scores by knowledge of treatment, can be examined by using the F statistic.

The F test is defined by the formula:

$$F = \frac{R_f^2 - R_r^2 \Big/ df_n}{1 - R_f^2 \Big/ df_d}$$

where R_f^2 is the amount of variance accounted for by the predictors in the full model, and R_r^2 is the amount of variance accounted for by the predictors in the restricted model.

The number of degrees of freedom associated with the numerator is the difference between the number of independent sources of information in the full model (unit vector and vectors for age, IQ, reading, pre-science knowledge and vectors representing group membership), and the independent sources in the restricted model (unit vector, age, IQ, reading, and previous science knowledge).

Thus: $$df_n = 6\text{-}5$$

The number of degrees of freedom associated with the denominator is defined as the difference between the number of subjects (N), and the number of independent sources of information in the full model. The degrees of freedom denominator varies with the size of the group studied. For the full group analysis N = 87 and

$$df_d = 87\text{-}6$$

Ten F ratios were calculated in each analysis performed. That is, the five hypotheses were examined using the STEP post-test results as the criterion to be predicted, and then re-examined using the teacher constructed post-test as the predictor criterion.

The hypotheses were generated from the comparisons among the sub-groups in the investigation. These sub-groups are presented in TABLE III.

The vectors used in the testing of the hypotheses are defined on page 40.

The hypotheses are shown on page 41. Full and restricted models are presented for each hypothesis.

TABLE III

SUBJECT SUB-GROUPS

Treatment	Sex		Total
	Girls	Boys	
Experimental	Cell 1	Cell 3	Cell 5
Control	Cell 2	Cell 4	Cell 6

VECTORS

X_1 .... Treatment

X_2 .... Pre-test science scores

X_3 .... Post-test scores STEP

X_4 .... Post-test scores Teacher test

X_5 .... Intelligence quotient

X_6 .... Pupil sex

X_7 .... Age

X_8 .... Reading comprehension score

X_9 .... All experimental subjects (treatment 1)

X_{10} .... All control subjects (treatment 2)

X_{11} .... Boys

X_{12} .... Girls

X_{13} .... $X_9 \times X_{11}$ Boys in experimental groups

X_{14} .... $X_9 \times X_{12}$ Girls in experimental groups

X_{15} .... $X_{10} \times X_{11}$ Boys in control groups

X_{16} .... $X_{10} \times X_{12}$ Girls in control groups

HYPOTHESES AND MODELS

Hypothesis

1 Knowledge of how students are grouped for instruction does not contribute significantly to the ability to predict science achievement scores between experimental subjects and control subjects, when controlling for pre-test science scores, reading comprehension scores, intelligence and age.

Full model:

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + a_9X_9 + a_{10}X_{10} + e$$

Restricted model:

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + e$$

1a Knowledge of how students are grouped for instruction does not contribute significantly to the ability to predict science achievement scores between experimental females and control females, when controlling for pre-test science scores, reading comprehension score, intelligence and age.

Full model:

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + a_{14}X_{14} + a_{16}X_{16} + e$$

Restricted model:

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + e$$

1b Knowledge of how students are grouped for instruction does not contribute significantly to the ability to predict science achievement scores between experimental males and control males, when controlling for pre-test science scores, reading comprehension scores, intelligence and age.

Full model:

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + a_{13}X_{13} + a_{15}X_{15} + e$$

Restricted model:

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + e$$

2 Knowledge of how students are grouped for instruction does not contribute significantly to the ability to predict science achievement scores between experimental females and experimental males, when controlling for pre-test science scores, reading comprehension scores, intelligence and age.

Full model:

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + a_{13}X_{13} + a_{14}X_{14} + e$$

Restricted model:

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + e$$

3 Knowledge of how students are grouped for instruction does not contribute significantly to the ability to predict science achievement scores between control females and control males, when controlling for pre-test science scores, reading comprehension scores, intelligence and age.

Full model:

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + a_{15}X_{15} + a_{16}X_{16} + e$$

Restricted model:

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + e$$

Models identical with these were used to test the hypotheses using the teacher constructed post-test (X_4) as the criterion measure.

CHAPTER IV

RESULTS OF THE INVESTIGATION

This chapter reports the results of the analysis of the data collected during the investigation. The null hypotheses are examined in terms of both criterion measures, the STEP post-test and the teacher constructed post-test. Since the group means on the post-tests are reported by the multiple linear regression programme as unadjusted means, it was felt necessary to generate adjusted means. These are reported in TABLE IV.

The group means on both criterion tests are influenced by the covariates of previous science knowledge, intelligence, reading comprehension, and age. These covariates are treated as a group in the study. It was felt that the isolation of the individual covariates would not contribute information of general educational usefulness since the covariates form a group of intercorrelated variables that, if used individually, would probably each contribute significantly to the variance.

The covariates above were chosen for inclusion in the study as they are representative of typical variables that need to be controlled for. To show the influence of the individual covariates, a stepwide regression analysis was performed on the covariates, and is reported following the analyses of the major part of the study.

TABLE IV

UNADJUSTED AND ADJUSTED MEANS
CRITERION POST TESTS

	STEP		POST*	
	Unadj.	Adj.	Unadj.	Adj.
Ex. boys	270.1	272.6	22.3	22.9
Ex. girls	265.5	262.9	22.1	22.7
Cont. boys	268.3	272.4	23.2	22.1
Cont. girls	264.8	264.2	21.2	21.1

* POST here refers to the teacher constructed post-test.

As can be seen from TABLE IV, the influence of the covariates did not result in large adjustments of the post tests' means. The adjusted means on the STEP test have an average difference of 2.4 points from the unadjusted means. On the teacher made test the average difference between adjusted and unadjusted means is 0.6.

RESULTS OF TESTING HYPOTHESES ON STEP CRITERION

Hypothesis 1

> Knowledge of how students are grouped for instruction does not contribute significantly to the ability to predict science achievement scores between experimental subjects and control subjects, when controlling for pre-test science scores, reading comprehension scores, intelligence and age.

The question to be examined is whether membership in the experimental or in the control group influences science achievement significantly.

A full model

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + a_9X_9 + a_{10}X_{10} + e$$

was constructed in which pre-test science scores (X_2), intelligence (X_5), age (X_7), and reading comprehension scores (X_8) were controlled for, and the group membership vectors (X_9,X_{10}) included.

The restricted model

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + e$$

assumed that group membership did not affect the subjects' science achievement, over and above that portion accounted for by the covariates.

The weights a_2,a_5,a_7,a_8,a_9, and a_{10} are unknowns which make the linear statements true (Kelly, p. 56).

The regression analysis yielded squared multiple correlations of 0.44 for the full model, and 0.43 for the restricted model.

When compared by means of an F test, the correlations produced an F ratio of 1.26. Since an F ratio greater than 3.96 would be required for significance at the 0.05 level (degrees of freedom numerator

1, and degrees of freedom denominator 81), the null hypothesis was not rejected.

The results of the analysis are shown in TABLE V.

Hypothesis 1a

> Knowledge of how students are grouped for instruction does not contribute significantly to the ability to predict science achievement scores between experimental females and control females, when controlling for pre-test science scores, reading comprehension scores, intelligence and age.

The hypothesis was tested by means of the same procedure as was used for hypothesis 1, since the question to be examined is whether membership in the experimental group or in the control group influences the science achievement of girls.

A full model

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + a_{14}X_{14} + a_{16}X_{16} + e$$

was constructed which included information about membership in the experimental or the control group (X_{14},X_{16}), as well as the following covariates: pre-test science scores (X_2), reading comprehension scores (X_8), age (X_7) and intelligence (X_5).

The restricted model

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + e$$

contained information only on the covariates. The effect of the comparison of the two models is to isolate group membership as the significant influence on the girls' science scores, over and above the variance accounted for by the covariates.

TABLE V

ANALYSIS OF COVARIANCE OF SCIENCE STEP TEST

FULL GROUP

HYPOTHESIS 1

	Adjusted means		R_f^2	R_r^2	Degrees of freedom			
	Exp.	Cont.			Num.	Denom.	F ratio	p
Hypothesis 1 (total group)	269.1	266.8	0.44	0.43	1	81	1.26	0.26
Hypothesis 1a (girls group)	265.5	264.2	0.48	0.43	1	39	3.64	0.06
Hypothesis 1b (boys group)	272.6	268.4	0.49	0.43	1	36	4.32	0.04*

*significant at .05 level

R_f^2 variance accounted for by full model

R_r^2 variance accounted for by restricted model

The weights a_2, a_5, a_7, a_8, a_{14}, a_{16}, make the linear statements true.

The squared multiple correlation for the full model was 0.48, and that for the restricted model 0.43. When subjected to an F test they yielded an F ratio of 3.64 (see TABLE V).

With 1 degree of freedom in the numerator, and 39 degrees of freedom in the denominator, the F ratio was not significant at the 0.05 level. The null hypothesis was therefore not rejected (TABLE V). Thus, girls in the experimental and control groups were found not to differ significantly in terms of their achievement on the STEP posttest.

Hypothesis 1b

> Knowledge of how students are grouped for instruction does not contribute significantly to the ability to predict science achievement scores between experimental males and control males, when controlling for pre-test science scores, reading comprehension scores, intelligence and age.

This hypothesis was tested by means of the same procedure used to test hypothesis 1a, as the question to be examined is whether membership in the experimental or in the control group influences the science achievement of boys.

A full model

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + a_{13}X_{13} + a_{15}X_{15} + e$$

was constructed which included information about the group membership, as well as the following covariates: science pre-test scores (X_2), reading comprehension scores (X_8), intelligence (X_5), and age (X_7).

The restricted model contained information only on the covariates.

Restricted model

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + e$$

Comparing the two models isolates group membership as the significant influence on science achievement, over and above the variance accounted for by the covariates.

The squared multiple correlation for the full model was 0.49, and for the restricted model 0.43. When subjected to an F test they yielded an F ratio of 4.32.

With 1 degree of freedom in the numerator, and 36 degrees of freedom in the denominator, the F ratio was significant at the 0.05 level. The null hypothesis was therefore rejected (TABLE V).

Thus, boys in the experimental and control groups were found to differ significantly in terms of their achievement on the STEP post-test.

Hypothesis 2

> Knowledge of how students are grouped for instruction does not contribute significantly to the ability to predict science achievement scores between experimental females and experimental males, when controlling for pre-test science scores, reading comprehension scores, intelligence and age.

This hypothesis recognizes that sex differences within the experimental group might influence the ability to predict science achievement scores even when the treatment for both sexes is the same.

The analysis again involves the multiple regression procedure

used to test the previous hypotheses. The group examined are those students given instruction in groups consisting of all boys or all girls. Full and restricted models were constructed as follows:

Full model

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + a_{13}X_{13} + a_{14}X_{14} + e$$

Restricted model

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + e$$

The restricted model assumes that no additional information over and above that attributed to the covariates (X_2,X_5,X_7,X_8) will be gained by knowing that the subjects are boys or girls if they are instructed in separate groups which are all boy or all girl groups. These groups are represented by the vectors X_{13} and X_{14} in the full model.

TAVLE VI shows the results of the regression analysis. The F ratio of 3.26 which results from comparing the two models is not significant at the 0.05 level, and the hypothesis is not rejected.

Hypothesis 3

> Knowledge of how students are grouped for instruction does not contribute significantly to the ability to predict science achievement scores between control females and control males, when controlling for science pre-test scores, reading comprehension scores, intelligence and age.

Full and restricted models were constructed as follows, and subjected to a multiple linear regression analysis.

TABLE VI

ANALYSIS OF COVARIANCE OF SCIENCE STEP TEST

FULL GROUP

HYPOTHESES 2 AND 3

	Adjusted means		R_f^2	R_r^2	Degrees of freedom		F ratio	p
	males	females			Num.	Denom		
Hypothesis 2 Ex. Group	272.6	265.5	0.48	0.43	1	35	3.26	0.07
Hypothesis 3 Cont. Group	268.4	264.2	0.45	0.43	1	40	1.53	0.22

R_f^2 variance accounted for by full model

R_r^2 variance accounted for by restricted model

Full model

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + a_{15}X_{15} + a_{16}X_{16} + e$$

Restricted model

$$X_3 = A_oU + a_2X_2 + a_5X_5 + a_7X_7 + a_8X_8 + e$$

The restricted model assumes that no additional significant information beyond that provided by the covariates is available if it is known whether subjects in mixed groups are boys or girls.

The squared multiple correlation coefficients for the full and restricted models were 0.45 and 0.43 respectively. When subjected to an F test they yield an F ratio of 1.53, much below that needed for 0.05 level significance, with 1 degree of freedom numerator and 40 degrees of freedom denominator. The probability that the difference occurs by chance exceeds 22% (see TABLE VI).

Thus, hypothesis 3 is not rejected, and no significant differences on STEP post-test results were found between boys and girls in the control group.

The models used to test the hypotheses, using STEP results as the criterion, are to be found in TABLE XXIV in Appendix D. Calculated constants and regression weights are included.

RESULTS OF TESTING HYPOTHESES ON TEACHER TEST

Since the multiple regression analysis was performed using two criterion measures, namely, STEP post-test results and teacher constructed post-test results, the hypotheses to be tested can be subjected to similar discussions for both criterion measures. Thus the analysis which follows is a replication of that described earlier with the exception that the criterion measure is X_4 (teacher constructed post-test scores) and not X_3 (STEP post-test scores).

The full and restricted models for each hypothesis are the same as those in the STEP criterion analysis, so that the results can be readily tabulated. TABLE VII reports the findings with respect to hypotheses 1, 1a, and 1b. TABLE VIII reports the findings for hypotheses 2 and 3.

The hypotheses tested, and a discussion of the models used to test them are to be found in the analysis for the STEP criterion, pages 46 to 53. TABLE XXIV in Appendix D shows the models used to test the hypotheses, along with the calculated constants and regression weights.

Both hypotheses 1a and 1b were not rejected since the F ratios in both instances were smaller than those required for the difference between the adjusted means for the groups on the teacher test to be significant at the 0.05 level.

The failure to reject the hypothesis 1a means that no significant difference was found between the adjusted mean scores for the experimental and control girls on the teacher constructed post-test.

TABLE VII

ANALYSIS OF COVERIANCE OF TEACHER POST TEST

FULL GROUP

HYPOTHESIS 1

	Adjusted means		R_f^2	R_r^2	Degrees of freedom			
	Exp.	Cont.			Num.	Denom.	F ratio	p
Hypothesis 1 (total group)	22.8	21.6	0.47	0.44	1	81	4.59	0.04*
Hypothesis 1a (girls group)	22.7	21.1	0.47	0.44	1	39	2.40	0.13
Hypothesis 1b (boys group)	22.9	22.1	0.45	0.44	1	36	0.88	0.35

*Significant at .05 level

R_f^2 variance accounted for by full model

R_r^2 variance accounted for by restricted model

TABLE VIII

ANALYSIS OF COVARIANCE OF TEACHER POST TEST

FULL GROUP

HYPOTHESES 2 AND 3

	Adjusted means		R_f^2	R_r^2	Degrees of freedom		F ratio	p
	males	females			Num.	Denom.		
Hypothesis 2 Ex. group	22.9	22.7	0.47	0.44	1	35	2.06	0.15
Hypothesis 3 Cont. group	22.1	21.1	0.48	0.44	1	40	3.11	0.08

R_f^2 variance accounted for by full model

R_f^2 variance accounted for by restricted model

In terms of hypothesis 1b, no significant difference between the adjusted mean scores for experimental boys and control boys was found.

Hypothesis 1, however, is rejected since the F ratio of 4.59 is greater than the critical value required for the difference between the adjusted mean scores for the groups to be significant at the 0.05 level.

The hypothesis was tested using a full model containing co-variate information (pre-test science scores, reading comprehension scores, intelligence and age) and X_9 and X_{10} which are group vectors, X_9 being the full experimental group, and X_{10} the full control group. The restricted model contains information on only the covariates.

Knowledge of which group the subjects belong to, experimental or control, permits a statistically better prediction of teacher post-test scores than lack of this knowledge.

It is interesting to note that while comparison of the experimental boys group with the control boys group, and of experimental girls group with the control girls group yields differences which are not significant, when the total experimental group (adjusted mean scores) is compared with the total control group, the difference in performance on the teacher constructed post-test is statistically significant at the 0.05 level.

TABLE XVIII presents a tabulation of the findings of the regression analyses for hypotheses 2 and 3. It can be readily seen that the F ratios are smaller than those required for rejection of

the hypotheses.

Thus it can be said that no significant difference was found on the teacher post-test between the performance of experimental girls and experimental boys. Also, no significant difference was found between the performance of control boys and control girls on the test.

TABLE IX summarizes the findings of the regression analysis for the hypotheses tested on both criterion measures. The models used in the regression analysis are presented in TABLE XXIV, Appendix D, together with calculated constants and regression weights.

TABLE IX

SUMMARY OF RESULTS OF TESTING HYPOTHESES ON TWO CRITERION POST TESTS

Hypothesis	STEP post-test	Teacher post-test
1	not rejected	rejected*
1a	not rejected	not rejected
1b	rejected*	not rejected
2	not rejected	not rejected
3	not rejected	not rejected

* significant at 0.05 level

OTHER FINDINGS

An examination of the unadjusted pre-test and post-test means on the STEP tests indicated that further analysis might be necessary.

While it had not been intended to use the STEP results to compare pre and post test scores, in view of the fact that the mean scores on the girls' post test results fell while the score for both boys groups rose, it was decided to determine if relevant information was being hidden in the full group analysis.

The means on the pre and post STEP tests are reported in TABLE X. Unadjusted means are shown.

TABLE X

UNADJUSTED MEANS PRE AND POST STEP

	Pre-STEP	Post-STEP	Difference
Experimental boys	264.4	270.1	+ 5.7
Experimental girls	264.3	262.9	- 1.4
Control boys	271.1	272.4	+ 1.3
Control girls	266.4	264.8	- 1.6

Multiple regression analysis was performed to test the hypotheses for the subjects of the two teachers separately. The hypotheses tested are those examined in the full group analysis (pages 46 to 59). TABLE XI shows the distribution of students according to teacher and instructional grouping.

TABLE XI

DISTRIBUTION OF STUDENTS BY TEACHER AND TREATMENT

	Teacher 1		Teacher 2	
	Boys	Girls	Boys	Girls
Experimental	10	15	10	6
Control	10	11	12	13
Total	20	26	22	19

RESULTS OF TESTING HYPOTHESES - TEACHER ONE

Since the experimental and control classes for each teacher are examined separately this further analysis does, in effect, introduce a variable of teacher performance into this part of the study. TABLES XII and XIII which follow show the results of the regression analysis of the data for the STEP criterion for the subjects taught by Teacher one. The models used to test the hypotheses are reported in TABLE XXV, APPENDIX D. Calculated regression weights are included.

TABLE XII

ANALYSIS OF COVARIANCE OF SCIENCE STEP TEST

TEACHER ONE

HYPOTHESIS 1

	Adjusted means		R_f^2	R_r^2	Degrees of freedom		F ratio	p
	Males	Females			Num.	Denom.		
Hypothesis 1 (total group)	270.1	265.7	0.51	0.48	1	40	2.41	0.12
Hypothesis 1a (girls group)	267.4	264.9	0.51	0.48	1	20	1.28	0.27
Hypothesis 1b (boys group)	272.9	266.6	0.54	0.48	1	14	1.91	0.18

R_f^2 variance accounted for by full model

R_r^2 variance accounted for by restricted model

TABLE XIII

ANALYSIS OF COVARIANCE OF SCIENCE STEP TEST

TEACHER ONE

HYPOTHESES 2 AND 3

	Adjusted means		R_f^2	R_r^2	Degrees of freedom			
	Males	Females			Num.	Denom.	F ratio	p
Hypothesis 2 Ex. group	272.9	267.4	0.54	0.48	1	19	2.79	0.11
Hypothesis 3 Cont. group	266.6	264.9	0.51	0.48	1	15	1.02	0.32

R_f^2 variance accounted for by full model

R_r^2 variance accounted for by restricted model

Hypothesis 1 postulates that no difference in science achievement will result between subjects grouped for instruction in single-sex groups or mixed sex groups (p.41), over and above that accounted for by the covariates of pre-test science scores, reading comprehension scores, intelligence and age.

No significant difference was found between the full experimental and full control groups for teacher one. Thus, hypothesis 1 is not rejected.

Hypothesis 1a examines the effect of instructing girls in all-girl groups and in mixed groups (p.41). Since the F ratio does not approach the magnitude required to reject the hypothesis, it must be concluded that knowledge of group membership does not significantly enhance the ability to predict science STEP scores, over and above the proportion of variance accounted for by the covariates. That is, the girls groups do not differ significantly as a result of the treatment.

Hypothesis 1b (p.42) is not rejected since the F ratio is smaller than that required to justify rejection at the 0.05 level of confidence. Thus it must be concluded that knowledge of which group the boys are in for instruction does not materially help in predicting STEP science scores, over and above that proportion of the variance accounted for by the covariates. The different treatments did not alter the boys' scores on the STEP test.

TABLE XIII shows the results of testing hypotheses 2 and 3 (p.42). Neither hypothesis is rejected, indicating no significant

differences between experimental boys and girls, nor any significant differences between control boys and control girls on the STEP criterion.

TABLE XIV reports the results of testing hypotheses 1, 1a, and 1b with the teacher constructed post test used as the criterion. The models used to test the hypotheses are reported in TABLE XXV, APPENDIX D, together with calculated regression weights.

While no differences were found between experimental and control boys on this criterion, the two groups of girls did differ significantly, over and above that proportion of variance accounted for by the covariates pre-test science scores, reading comprehension scores, intelligence and age. In addition, the difference between the performance of the full experimental and full control groups on the teacher constructed post-test is statistically significant. Consequently, hypotheses 1 and 1a are rejected, while hypothesis 1b is not rejected (see TABLE XVI).

When differences between boys and girls in the experimental group, and differences between boys and girls in the control groups are examined (hypotheses 2 and 3, page 42), significant differences were found in both instances. TABLE XV shows the results of the analyses.

Of the hypotheses tested with the results obtained by teacher one, none are rejected when the STEP test is used as the criterion measure. However, if the teacher constructed post-test is used as the criterion, hypotheses 1, 1a, 2 and 3 are rejected (see TABLE XVI).

TABLE XIV

ANALYSIS OF COVARIANCE OF TEACHER POST TEST

TEACHER ONE

HYPOTHESIS 1

	Adjusted means		R_f^2	R_r^2	Degrees of freedom			
	Exp.	Cont.			Num.	Denom.	F ratio	p
Hypothesis 1 (total group)	23.7	21.3	0.53	0.43	1	40	9.12	0.004**
Hypothesis 1a (girls group)	23.4	20.4	0.54	0.43	1	20	5.19	0.03*
Hypothesis 1b (boys group)	23.9	22.2	0.47	0.43	1	14	1.08	0.31

R_f^2 variance accounted for by full model

R_r^2 variance accounted for by restricted model

* significant at 0.05 level

** significant at 0.01 level

TABLE XV

ANALYSIS OF COVARIANCE OF TEACHER POST TEST

TEACHER ONE

HYPOTHESES 2 AND 3

	Adjusted means		R_f^2	R_r^2	Degrees of freedom		F ratio	p
	Males	Females			Num.	Denom.		
Hypothesis 2 Ex. group	23.9	23.4	0.54	0.43	1	29	4.63	0.04*
Hypothesis 3 Cont. group	22.2	20.4	0.56	0.43	1	15	4.95	0.04*

R_f^2 variance accounted for by full model

R_r^2 variance accounted for by restricted model

* significant at 0.05 level

TABLE XVI

SUMMARY OF RESULTS OF HYPOTHESES ON TWO CRITERION POST TEST FOR TEACHER ONE

Hypothesis	STEP post-test	Teacher post-test
1	not rejected	rejected**
1a	not rejected	rejected*
1b	not rejected	not rejected
2	not rejected	rejected*
3	not rejected	rejected*

** significant at 0.01 level

* significant at 0.05 level

RESULTS OF TESTING HYPOTHESES - TEACHER TWO

The analyses of the hypotheses for Teacher two are now presented. Since the hypotheses to be tested are the same as those tested previously in the full analysis (pages 46 to 59), the models used in the regression analysis are the same as those used before. These models with their calculated regression weights are to be found in TABLE XXVI in APPENDIX D.

Two tables follow which report the results of the analyses for Teacher two using the STEP criterion.

None of the hypotheses were rejected. The F ratios are such that no significant differences were found between treatment groups, or between boys and girls subjected to similar treatments (TABLES XVII and XVIII).

TABLE XVII

ANALYSIS OF COVARIANCE OF SCIENCE STEP TEST

TEACHER TWO

HYPOTHESIS ONE

	Adjusted means		R_f^2	R_r^2	Degrees of freedom		F ratio	p
	Exp.	Cont.			Num.	Denom.		
Hypothesis 1 (total group)	265.7	266.9	0.51	0.51	1	35	0.0	1.00
Hypothesis 1a (girls group)	259.1	265.6	0.59	0.51	1	13	2.58	0.13
Hypothesis 1b (boys group)	272.4	268.3	0.58	0.51	1	16	2.74	0.11

R_f^2 variance accounted for by full model

R_r^2 variance accounted for by restricted model

TABLE XVIII

ANALYSIS OF COVARIANCE OF SCIENCE STEP TEST

TEACHER TWO

HYPOTHESES 2 AND 3

	Adjusted means		R_f^2	R_r^2	Degrees of freedom		F ratio	p
	Males	Females			Num.	Denom		
Hypothesis 2 (Exp. group)	272.4	259.1	0.59	0.51	1	10	2.18	0.16
Hypothesis 3 (Cont. group)	268.3	265.6	0.52	0.51	1	19	0.20	0.65

R_f^2 variance accounted for by full model

R_r^2 variance accounted for by restricted model

The results of the analyses using the teacher constructed test as the criterion with subjects of teacher two are presented in TABLES XX and XXI.

As can be seen from the tables, knowledge of group membership does not enhance the ability to predict, over and above the proportion of variance accounted for by the covariates, scores on the teacher constructed post-test.

The models used to test the hypotheses can be found in TABLE XXVI in APPENDIX D. The hypotheses tested are the same as those tested previously in the full analysis (pages 46 to 59).

TABLE XIX summarizes the findings of the regression analyses for the hypotheses for both criterion measures for Teacher two.

TABLE XIX

SUMMARY OF RESULTS TEACHER TWO

Hypothesis	STEP post-test	Teacher post-test
1	not rejected	not rejected
1a	not rejected	not rejected
1b	not rejected	not rejected
2	not rejected	not rejected
3	not rejected	not rejected

TABLE XX

ANALYSIS OF COVARIANCE OF TEACHER POST TEST

TEACHER TWO

HYPOTHESIS 1

	Adjusted means		R_f^2	R_r^2	Degrees of freedom			
	Exp.	Cont.			Num.	Denom.	F ratio	p
Hypothesis 1 (total group)	21.5	21.7	0.52	0.51	1	35	0.01	0.91
Hypothesis 1a (girls group)	20.8	21.8	0.53	0.51	1	13	2.59	0.13
Hypothesis 1b (boys group)	22.1	21.5	0.52	0.51	1	16	2.74	0.11

R_f^2 variance accounted for by full model

R_r^2 variance accounted for by restricted model

TABLE XXI

ANALYSIS OF COVARIANCE OF TEACHER POST TEST

TEACHER TWO

HYPOTHESES 2 AND 3

	Adjusted means		R_f^2	R_r^2	Degrees of freedom			
	Males	Females			Num.	Denom	F. ratio	p
Hypothesis 2 (Ex. group)	22.1	20.8	0.53	0.51	1	10	0.31	0.58
Hypothesis 3 (Cont. group)	21.5	21.8	0.52	0.51	1	19	0.05	0.83

R_f^2 variance accounted for by full model

R_r^2 variance accounted for by restricted model

SUMMARY OF THE RESULTS

The results of the data analysis may be summarized as follows:

For the full group analysis using both criterion measures, and with the covariates of previous science knowledge, reading comprehension ability, intelligence and age, controlled for the findings indicate:

1 No significant difference was found in the adjusted mean performance between the full experimental and full control groups on the STEP criterion. However, the experimental group was superior to the control group on the criterion of the teacher constructed post-test.

1a There was no significant difference in the adjusted mean performance between girls instructed in all-girl groups, and girls instructed in mixed groups on the STEP post-test or on the teacher constructed post-test.

1b On the criterion of STEP post-test the experimental boys were significantly superior to the control boys in adjusted mean performance. No significant difference was found between the mean performance of the groups on the teacher constructed post-test.

2 No significant differences were observed between experimental males and experimental females on either criterion measure.

3 No significant differences were observed between control males and control females on either of the criterion measures.

The data were analysed for the two teachers in the study separately. Using the covariates of previous science knowledge, reading comprehension ability, intelligence and age to control for student variability, the following results were obtained for the subjects of Teacher One.

1 No significant difference was found in the adjusted mean performance of the full experimental and full control groups on the STEP criterion. However, on the teacher post-test criterion the adjusted mean performance of the experimental group was significantly superior to that of the control group.

1a No significant differences in adjusted mean performance between girls in all-girl groups and girls in mixed groups were found when using the STEP criterion.

When using the teacher constructed post-test as the criterion measure, the adjusted mean performance of the experimental girls was superior to that of the control girls.

1b No significant differences in adjusted mean performance were observed between boys instructed in all-boy groups and boys instructed in mixed groups on either of the criterion measures.

2 No difference in adjusted means scores of a significant nature was found between experimental females and experimental males when using the STEP criterion.

However, on the teacher post-test criterion the adjusted mean performance of the experimental group was significantly superior to that of the control group.

3 Similar results were obtained for this hypothesis as for hypothesis 2. Males mean performance was superior to that of the females on the teacher post-test criterion.

For Teacher Two no significant differences were observed between the adjusted mean scores for the groups for the different treatments, for any of the hypotheses on either of the two criterion measures, over and above that proportion of the variance accounted for by the covariates.

STEP-WISE REGRESSION ANALYSIS

The covariates included in the major part of this study were considered as a group of variables, representative of the kind of variables which influence achievement, and which need to be controlled for in studies such as this. The fact that the individual covariates' contribution to the prediction of criterion scores differs, was not considered to be central to this study. However, it was decided to examine the extent to which the individual covariates contribute to the prediction of the criterion scores. Thus, a step-wise regression analysis was performed on the covariates for both of the criteria used.

The step-wise regression procedure used permits each covariate to enter the analysis separately. The degree of contribution to the prediction of the criterion is calculated, and the significance of this contribution examined by using the F statistic. The per cent of the variance accounted for by the covariates is reported at each step.

TABLE XXII reports the results of the step-wise regression analysis for the covariates on the STEP criterion. The regression equation with the numerical coefficients included follows:

$$X_5 = 69.75 + .31X_1 + .39X_2 + .34X_3 + .32X_4$$

X_1 is reading comprehension, X_2 intelligence, X_3 previous science knowledge and X_4 age. As can be seen from TABLE XXII, the covariates reading comprehension, intelligence and previous science

TABLE XXII

STEP-WISE REGRESSION ANALYSIS OF FOUR COVARIATES FOR

STEP POST TEST

	F ratio	p	% of variance
Reading score	38.56	0.00	31.20
IQ	8.22	0.01	37.33
Pre-test score	5.22	0.02	41.04
Age	3.53	0.06	43.48

The values of F, p, and % of variance are the TOTAL values at each step.

knowledge all contribute significantly to the prediction of STEP scores. Age is not a significant predictor of STEP scores, although the F ratio at this step does approach significance.

TABLE XXIII reports the results of the analysis using the teacher test as the criterion. The regression equation and coefficients follow. In comparing the two sets of results it should be noted that the covariates have different designations in each analysis.

$X_5 = -19.54 + .12X_1 + .09X_2 - .03X_3 + .04X_4$

X_1 is previous science knowledge, X_2 intelligence, X_3 age and X_4 reading comprehension. As TABLE XXIII shows, only the covariates previous science knowledge and intelligence are significant contributors to the prediction of scores on the teacher post-test.

It is interesting to note that the two significant predictors on the teacher post-test account for a greater amount of the variance than the three significant covariates on the STEP post-test.

While age is not a significant covariate on either of the criterion tests, reading comprehension is significant with respect to the STEP criterion, but not the teacher test criterion. It is possible that this results because the STEP test is harder to read than the teacher constructed test, and is thus more a test of reading ability than the teacher made test.

TABLE XXIII

STEP-WISE REGRESSION ANALYSIS OF FOUR COVARIATES FOR

TEACHER POST-TEST

	F ratio	p	% of variance
IQ	44.40	0.00	34.31
Pre-test scores	12.82	0.01	43.01
Reading scores	1.10	0.29	43.75
Age	0.28	0.59	43.94

The values of F, p, and % of variance are TOTAL values at each step.

CHAPTER V

SUMMARY, CONCLUSIONS, IMPLICATIONS AND SUGGESTIONS FOR FURTHER RESEARCH

SUMMARY

The purpose of this study was to determine if grade six students grouped in all-boy and all-girl groups for instruction in science would achieve at a significantly higher level than those taught in mixed groups.

Four grade six classes were involved in the study. Two classes were designated by chance as experimental. In these classes the students were grouped in all-boy and all-girl groups. Two male teachers participated in the study, each one instructing both an experimental and a control class.

Data were collected for previous science knowledge, reading comprehension ability, intelligence and age. These were included in the analysis of the post-test results to control for student variability. The total testing procedure was carried out by the classroom teachers.

The post science achievement tests were a STEP multiple choice test, and a multiple choice test made up by the investigator and the participating teachers. The hypotheses were tested using the two criterion tests separately.

In the light of the comparison of the pre and post-STEP scores it was decided to re-examine the hypotheses for the students of each

teacher as separate groups. Thus, an analysis similar to the full group analysis was performed on the data obtained from the two classes taught by teacher one, and repeated for the data obtained from the two classes taught by teacher two.

An examination of the contribution of the individual covariates to the prediction of science achievement was also performed.

Conclusions With Respect to Hypotheses: Full Group

Hypothesis 1: The results of testing hypothesis one were inconclusive. When tested using the STEP test criterion, the hypothesis was not rejected. In other words, no improvement in science achievement scores at a significant level resulted from instruction being given to all-boy and all-girl groups. However, when tested using the teacher test criterion, the hypothesis was rejected. Thus, the experimental subjects scored significantly better on the one criterion test, but not on the other.

Hypothesis 1a: When the results for the girls in the study were examined on both criteria, they indicated that no improvement in science scores was attained by instructing the girls in all-girl groups in science.

Hypothesis 1b: When the results for the boys in the study were examined, they were again inconclusive. Boys in the experimental groups achieved significantly more than boys in the control groups on the STEP criterion. When using the teacher test criterion measure, no significant differences between the achievement of experimental boys

and the achievement of control boys were observed.

Hypothesis 2: No significant differences in science achievement were found between boys and girls in the experimental groups on either of the criterion measures. Thus, the sex of the subject did not influence science achievement significantly among the experimental subjects.

Hypothesis 3: Similar results as for hypothesis two were found when examining differences in achievement between boys and girls in the control groups. In other words, no significant differences attributable to the sex of the subject were found in science achievement.

From the examination of the hypotheses for the full group, one would generally conclude that grouping by sex for instruction in science at the grade six level does not significantly enhance science achievement. The fact that some contradictory evidence appears in the analysis could be explained by the difference in validity of the two post tests, the teacher constructed test being more closely related to the content concepts of the teaching unit. Also, the STEP test claims a high degree of reliability, while the teacher test was found to be reliable at the 0.52 level.

An interesting factor noted in TABLE X, that of the apparent regression of the girls' scores on the STEP tests, led to a decision to repeat the analysis of the data for each teacher separately. While it had not been intended to compare pre and post-STEP results, when the unadjusted means were examined it appeared that further analysis

might disclose information hidden in the full group analysis.

Conclusions With Respect to Hypotheses: Teacher One

None of the hypotheses were rejected when examined using the STEP test as the criterion measure. In effect, grouping students by sex for instruction in science at the sixth grade level apparently does not significantly influence scores on the STEP science test.

It is interesting to note, however, that when the hypotheses are examined with the teacher post-test used as the criterion, different results were observed.

Hypothesis 1: The total experimental group's science achievement was significantly higher than the total control group's, indicating that grouping for instruction by sex improved the subjects' score on the teacher post-test.

Hypothesis 1a: The girls' results were similar to the full group's. On the teacher post-test, the girls in the experimental groups scored significantly higher than the girls in the control groups. For the girls in these groups, science achievement on the teacher post-test was enhanced when instruction was given to all-girl groups.

Hypothesis 1b: This hypothesis was not rejected, showing that grouping boys in all-boy groups did not improve their achievement on the teacher post-test.

Hypothesis 2: This hypothesis was rejected. Thus, differences in achievement which were significant were found between the boys and girls in the experimental group.

Hypothesis 3: This hypothesis was also rejected. The boys in the control group also out-performed the girls in the control group on the teacher post-test. Since both hypothesis two and hypothesis three were rejected, it must be concluded that sex differences in achievement are more significant than differences resulting from grouping practices, when the teacher test is used as the criterion measure.

The results of testing the hypotheses using the data obtained from the subjects of teacher one, are then inconclusive. It appears that when achievement is measured using the STEP test as criterion, no significant differences are observed. However, the evidence indicates that differences in achievement as measured by the teacher constructed post-test, do apparently exist.

Conclusions With Respect to Hypotheses: Teacher Two

None of the hypotheses were rejected when examined using either the STEP criterion, or the teacher post-test criterion.

None of the sub-groups instructed by teacher two showed any significant differences in science achievement from any other group. Thus, treatment received had no significant influence on science achievement scores, nor did the sex of the subjects in the groups.

The fact that the results for the subjects of teacher one differed from the results for the subjects of teacher two might be ascribed to differences in the mode of instruction. It is quite possible that more emphasis was placed on group activities in the room taught by teacher one, and that this led to superior performance on

the teacher test by the students in this room. However, it must be recognized that unknown factors may have influenced the results.

ANALYSIS OF THE CONTRIBUTION OF THE COVARIATES

The examination of the covariates (previous science knowledge, reading comprehension ability, intelligence and age) revealed that only two of the covariates contributed significantly to the prediction of both the STEP post-test scores, and the teacher constructed post-test scores. These were previous science knowledge and intelligence.

Reading ability contributed significantly to the prediction of STEP post-test scores only, indicating that the STEP test might require greater proficiency in reading than the teacher post-test. This factor might also contribute to the inconclusiveness of the findings with respect to the subjects of teacher one.

IMPLICATIONS

Since little conclusive evidence was found to support the general hypothesis that grouping by sex for instruction in science enhances achievement in science, it would appear that the establishment of class groups using criteria other than that of the sex of the subject is advisable.

Previous science knowledge and intelligence were both significant predictors of science achievement, and it is possible that these attributes could be the basis for establishing instructional groups in elementary school science.

The inconclusive nature of the results of this study tend to follow those of other grouping studies cited (pages 22-23). No small group studies were found in the research of the literature, but Wisenthal (1965), Fahrner and Cronin (1963), and Scott (1964) report results indicating some interaction between sex and science achievement in single-sex classes and in single-sex schools. No distinct patterns of significant differences in achievement in science emerged from these studies.

The question arises as to the advisability of grouping for instruction so as to develop the strengths possessed by the students, or to group so as to remediate their weaknesses. No simple answer can be given, but it would appear that thought should be given to student needs when arranging groups, and that this thought should encompass significant predictors of achievement in science as well as student interest in the subject.

The differences in results noted when testing the hypotheses for the two teachers separately might lead one to conclude that some teachers could enhance their students' science achievement by using sex as a grouping criterion. However, more investigation in this area would be necessary.

Garai and Scheinfeld (1968) suggest that boys would score higher in science than girls would if both were taught in single-sex classes. They suggest also that a masculine approach to teaching science might contribute to feminine failure to do well. While the boys in this study did score higher than the girls for all sub-groups

examined (TABLE IV, page 52), only the results reported for teacher one on the teacher made test were consistently significant (TABLE XIV, page 66). Thus, the results of this study do not generally support the assertion that a masculine approach affects the girls' results, since no significant differences were found for the subjects of teacher two on either criterion.

The programme of instruction in elementary school science insists that student involvement in the 'process' of science is essential, and that this can be achieved only if the students participate directly in concrete science experiences. The logistics of classroom experimentation in science suggests that small group activities are necessary for the implementation of the programme, and which criteria are to be used in the formation of such groups remains an important question.

SUGGESTIONS FOR FURTHER RESEARCH

A number of areas of further research are apparent. The study was carried out only in grade six. Possibly, different results might be obtained with grouping by sex if it was implemented earlier in the elementary school science programme.

No attempt was made to find out if upper achievers in science would benefit more from grouping by sex than achievers at other levels. Categorization of students into average, high and low achievers in science would permit an examination of these aspects. Experimental and control groups could be established at each level, and the inter-

action between achievement level and grouping treatment examined.

One of the criteria for choosing the BIOGEOGRAPHY unit was that it contained material of interest to both boys and girls. It is possible that a unit of instruction more directly of interest to boys would enable boys instructed in all boy groups to achieve more than boys instructed in mixed groups. Similarly, instructional material of greater interest to girls might result in higher science achievement for girls in all girl groups.

The teachers in the study were both males. It would be interesting to investigate the interaction effects of both men and women teachers with subjects of both sexes, in experimental and control grouping situations.

Finally, science achievement cannot be measured simply by using multiple choice tests. Over a period of time longer than the eight weeks of the study, an evaluation instrument more comprehensive in nature could be designed. This would permit the evaluation of the full spectrum of skills, attitudes and concepts which constitute the elementary school science programme.

BIBLIOGRAPHY

Anderson, R. "Children's Ability to Formulate Mental Models to Explain Natural Phenomena," Journal of Research in Science Teaching, 3:326-332, 1965.

Atkin, J. "Science in the Elementary School," Review of Educational Research, 34:263-285, 1964.

Atkinson, C. "The Effect of Sex Differences on the Study of General Science," Journal of Educational Research, 24:61-66, 1931.

Barker, D. "An Attempt to Investigate the Effects of the Informal Use of a 'Discovery Table' on the Scientific Knowledge of Primary School Children," Educational Research, 7:155-158, 1965.

Becker, L. "An Analysis of the Science and Mathematics Achievement of Gifted Children Enrolled in Segregated, Partially Segregated, and Non-segregated Classes," Unpublished Thesis, Dissertation Abstracts, 1964.

Bennett, G., Seashore, H., & Wesman, A. "The Differential Aptitude Test: An Overview," Personal and Guidance Journal, 35, 81-91, 1956.

Bentzen, F. "Sex Ratios in Learning," National Elementary School Principal, 46:13-17, 1966.

Boutwell, W. "What's Happening in Education?" PTA Magazine 57:25-26, 1963.

Butts, D. "The Degree to which Children Conceptualize from Experiences," Journal of Research in Science Teaching, 1:135-143, 1963.

Cardon, B. "Sex Differences in School Achievement," Elementary School Journal, May 427-434, 1968.

Carey, G. "Sex Differences in Problem-solving Performance," Journal of Abnormal and Social Psychology, 56:256-260, 1958.

Clark, B. "An Experiment in Cultivating Creative Thinking Abilities in the Classroom," Dissertation Abstract 29, 1158A,4,1968.

Clark, W. & Willis, W. "Boys and Girls - Are There Significant Ability and Achievement Differences?" Phi Delta Kappan, Nov. 73-76, 1959.

Cooley, W. "Current Research on the Career Development of Scientists," Journal of Counselling Psychologists, 11, 88-93, 1964.

Cropley, A., & Field, T. "Achievement in Science and Intellectual Style," Journal of Applied Psychology, 53:132-135, 1969.

Curriculum Guide for Elementary Schools, 1969, Department of Education, Edmonton, Alberta.

DeLorenzo, L. "A Comparison of Science Achievement of 6th Grade Pupils by Regular Classroom Teachers," Science Education, 47:202-205, 1963.

DeRoche, E. "A Study of the Effectiveness of Selected Creative Exercises on Creative Thinking and the Mastery of a Unit in Elementary Science," Dissertation Abstracts, 27:4162A, 1966.

Dunfee, M. Elementary School Science: A Guide to Current Research, Association for Supervision and Curriculum Development, N.E.A., Washington, 1967.

Durante, R. "An Experiment With Single Sex Classes in Port Alberni," Principal's Journal, 7:25-29, 1967.

Epstein, J. & Wilson, R. "Sex Differences in the Elementary School," National Elementary School Principal, 46:8-12, 1966.

Farhner, J., & Cronin, J. "Grouping by Sex," N.E.A. Journal, 52:16-17, 1963.

Ferguson, G. Statistical Analysis in Psychology and Education, New York: McGraw-Hill Book Company, 1966.

Fogelman, K. "Piagetian Tests and Sex Differences," Educational Research (Britain), 12:154-155, 1970.

Fox, D. The Research Process in Education, New York: Holt, Rinehart, and Winston, 1969.

Garai, J., & Scheinfeld, B. "Sex Differences in Mental and Behavioural Traits," Genetic Psychology Monographs, 77:169-299, 1968.

Goldschmid, M. "Different Types of Conservation and Non-conservation and their Relation to Age, IQ, M.A., and Vocabulary," Child Development, 38:1229-1246, 1967.

Goodenough, F. "The Measurement of Mental Growth in Childhood," in Manual of Child Psychology, New York: John Wiley & Sons, 1954.

Herman, B., & Criscuolo, N. "Sex Grouping," Instructor, March 97-98, 1968.

Hill, K., & Dusek, J. "Children's Achievement Expectations as a Function of Social Reinforcement, Sex of Subject, and Test Anxiety," Child Development, 40:547-557, 1969.

Hillerich, R. "A Brief Study of Research on Interclass Grouping at the Elementary School Level," U.S. Department of Health, Education and Welfare, Office of Education, March, 1967.

Hurd, A. "Sex Differences in Achievement in Physical Sciences," Journal of Educational Psychology, 25:70, 1934.

Jones, P. "An Analysis of Sex Differences in Alberta School Population," Unpublished Doctoral Thesis, University of Alberta, Edmonton, 1969.

Kelly, F. et al. Multiple Regression Approach, Carbondale: Southern Illinois University Press, 1969.

Kostic, M. "A Study of Transfer: Sex Differences in the Learning Process," The Journal of Educational Psychology, 45:449-458, 1954.

Koutnic, P. "The Effect of Cognitive Style, Sex and Degree of Completeness of an Electrical Data Background in Student Prediction of Circuital Function," Unpublished Doctoral Thesis, University of Indiana, 1968.

Lamkin, F. "A Personality Variable of Preadolescent Youth in Relation to the Elementary School Program," Unpublished Doctoral Thesis, Virginia University, Charlottsville, 1967.

Lewis, D. "The Factorial Nature of Attainment in Elementary Science," British Journal of Educational Psychology, 34:1-8, 1964.

Lyles, T. "Grouping by Sex," National Elementary School Principal, 46:38-41, 1966.

Maccoby, E. (ed.) The Development of Sex Differences, Stanford, Stanford University Press, 1966.

Manual for Interpreting Scores - Science, Cooperative Test Division, Educational Testing Services, Princeton, N.J., 1957.

McFarland, W. "Are Girls Really Smarter?" Elementary School Journal, 70:14-19, 1969.

Morgenstren, A. Grouping in the Elementary School, New York; Pitman, 1966.

Parsley, K. & Powell, M. "Are There Really Sex Differences in Achievement?" The Journal of Educational Research, 57: 268-270, 1964.

Peltier, G. "Sex Differences in the School: Proposed Solution," Phi Delta Kappan, 46:182-185, 1966.

Perrodin, A. "Children's Attitude Towards Elementary School Science," Science Education, 50:214-218, 1966.

Ramsey, G., & Howe, R. "An Analysis of Research Related to Instructional Procedures in Elementary School Science," Science and Children, 7:25-36, 1969.

Review of Educational Research: Science and Mathematics, 39:377-413, 1969.

Richer, S., & Breton, R. "School Organization and Student Differences," Canadian Education Research Digest, 8:20-37, 1968.

Rowland, G. "Sex Differences in the Science Background of Elementary School Children," Science Education, 52:179-180,1968.

Scott, N. "The Relationship of Inductive Reasoning and Cognitive Styles in Categorization Behaviour to Science Concept Achievement in Elementary School Children," Unpublished Doctoral Thesis, Wayne University, 1962.

Scott, N. "Science Concept Achievement and Cognitive Functioning," Journal of Research in Science Teaching, 2:7-16, 1964.

Sheehan, M. "Problem Solving," The Journal of Experimental Education, 36:84-87, 1968.

Sigel, I. "Rationale for Separate Analyses of Male and Female Samples on Cognitive Tasks," Psychological Record, 15, 369-376, 1965.

Slinn, P. "Teacher Influence and Pupil Achievement, Unpublished Masters Thesis, University of Alberta, Edmonton, 1969.

Smock, C., & Holt, B. "Children's Reaction to Novelty: An Experimental Study of Curiosity Motivation," Child Development, 33:631-642, 1962.

Stanchfield, Jo. "Differences in Learning Patterns of Boys and Girls," Unpublished Doctoral Thesis, Dissertation Abstracts, 1966.

Stroud, J., & Lindquist, E. "Sex Differences in Achievement in the Elementary and Secondary School," The Journal of Educational Psychology, 33:657-667, 1942.

Swan, M. "An Exploratory Study of Science Achievement as it Relates to Science Curricula and Programs at the Sixth Grade Level in Montana Public Schools," Unpublished Doctoral Thesis, University of Missoula, 1965.

Terman, L., & Tyler, L. "Psychological Sex Differences," in Manual of Child Psychology, New York: John Wiley & Sons, 1954.

Tyler, R. "Resources, Models and Theory in the Improvement of Research in Science Education," Journal of Research in Science Teaching, 5:43-51, 1967.

Uhlhorn, K. "The Preparation, Use and Application of a Science Experience Inventory," Unpublished Doctoral Thesis, Iowa State University, 1963.

Waetjen, W., & Grambs, J. "Sex Differences: A Case of Educational Evasion," Teachers College Record, 65:261-271, 1963.

Waetjen, W. "Learning and Motivation: Implications for the Teaching of Science," Science Teacher, 32:22-26, 1965.

Walberg, H. "Dimensions of Scientific Interests of Boys and Girls," Science Education, 51:111-116, 1967.

Walker, H., & Lev, J. Elementary Statistical Methods, Toronto: Holt, Rinehart, and Winston, 1958.

Watson, F. "Teaching Science to the Average Pupil," The Science Teacher, 34:24-26, 1967.

Weschler, D. Sex Differences in Intelligence: The Measurement and Appraisal of Adult Intelligence, Williams and Wilkins, 1958.

Wisenthal, M. "Sex Differences in Attitudes and Attainment in Junior Schools," British Journal of Educational Psychology, 35: 79-85, 1965.

Yates, A. (ed.) Grouping in Education, New York: John Wiley & Sons, 1967.

APPENDIX A

SAMPLE TEST ITEMS FROM STEP SCIENCE

TEST FORM A

PART TWO

Pets In Class

Miss Damrin's class kept several pets in their classroom.

1 One afternoon they were looking at the beautiful, tiny, tropical fish in the aquarium. Jane asked, "Why is the water heated for these fish?" Which pupil gave the best answer?

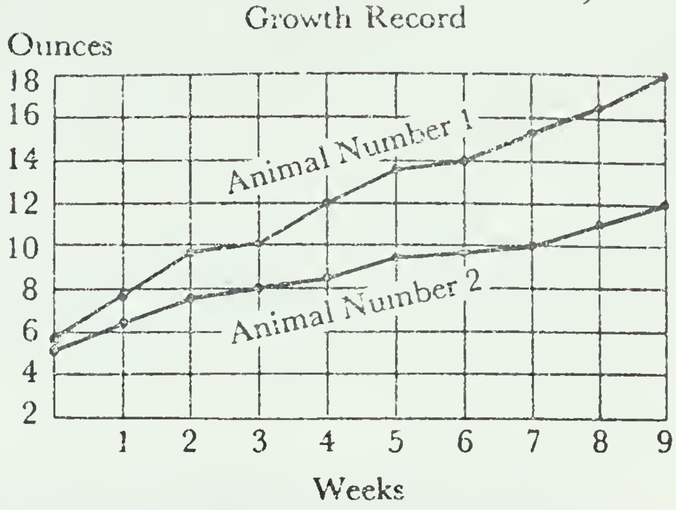
A Tom said, "To keep the water from freezing."
B Bill said, "To kill harmful germs which might be in the water."
C Ruth said, "To keep the water clean and clear for the fish."
D Bob said, "To give the fish the kind of home in which they can best live."

2 For their classroom, Ruth's father gave the children two cages and two little guinea pigs which were the same size. They wanted to find out if dark whole-wheat bread is better for health and growth than white bread. How could they best carry out their experiment?

E Give both guinea pigs white bread.
F Give both guinea pigs whole-wheat bread.
G Give one guinea pig whole-wheat bread and the other white bread.
H Give one guinea pig whole-wheat bread, and the other no bread at all.

3 Bill's father also gave the class a pair of guinea pigs. The children gave the two guinea pigs different diets. The children weighed their guinea pigs each Friday. They kept a growth record. This is their chart.

Growth Record



From this chart, can you tell what happened to the guinea pigs?

A Both gained the same amount of weight.
B Number 1 gained more than number 2.
C Number 2 gained more than number 1.
D Number 2 lost weight.

4 From this chart what was the difference in weight between the two guinea pigs after nine weeks?

E Number 1 weighed 6 ounces more than number 2.
F Number 1 weighed 18 ounces more than number 2.
G Number 2 weighed 6 ounces more than number 1.
H Number 2 weighed 12 ounces more than number 1.

5 Mary took one guinea pig home from school and fed him properly over the week end. The guinea pig lay quiet and rested on the bottom of the cage. Billy took the other one home. He forgot to feed him on Monday morning. The guinea pig was restless, ran around the cage, and bit Billy's finger. What do you think the class learned?

A Guinea pigs can live for days without food.
B Guinea pigs sleep when they are hungry.
C Guinea pigs fight each other for food.
D Hunger changes a guinea pig's actions.

Go on to the next page.

APPENDIX B

Sample test items from the post-test constructed by cooperating teachers and investigator:

James wants to prepare an environment for several water insects he caught in a nearby pond. His insects will be most comfortable if he uses

a. water from the tap
b. water from the pond
c. distilled (very pure) water

If you placed a slice of moistened bread in the freezer, another in the refrigerator, and a third in a warm oven, after several days where would you be most likely to find mold growing?

a. on the bread in the warm oven
b. on the bread in the freezer
c. on the bread in the refrigerator

John put a spoonful of yeast and a pint of water in each of three jars. He then added two tablespoons of sugar to one of the jars. In the second jar he put one teaspoon of sugar. The third jar had no sugar added. In which jar would you expect to see most increase in yeast population take place?

a. in the jar containing no sugar
b. in the jar containing one teaspoon of sugar
c. in the jar containing two tablespoons of sugar

The fact that many places in the world have a water shortage is not due to lack of water, but to

a. too many people using water
b. too many plants using water
c. an uneven rainfall over the world

APPENDIX C

BIOGEOGRAPHY ACTIVITIES

Big Ideas:

1. Environments

   Level 2

   Level 3

2. Environmental change

   Level 2

   Level 3

3. Diversity in life

   Level 2

   Level 3

4. Adaptation

   Level 2

   Level 3*

5. Time and life

   Level 2

   Level 3

6. Man and environment

   Level 2*

One sample activity from Big Idea 5 Level 3 follows.

*Home activities which were changed to classroom activities. All other activities are classroom activities.

Classroom Experiment

To prepare for this experiment, study the picture and list the materials you will need. Do this on your record page, under "Notes for the Classroom Experiment."

After you have studied the experiment carefully and made your notes, write your prediction on your record page, under "Predicting the Outcome of the Classroom Experiment." When you have completed this, you are ready to do the experiment. Write your observations and conclusions in the proper spaces under 3 and 4 of Part IV of your record page.

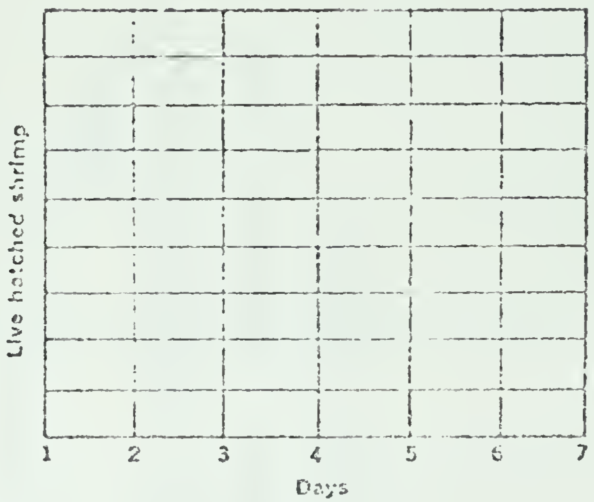
The Problem: How does the environment affect the growth rate of a population?

Directions:

Obtain dried brine shrimp eggs from a tropical fish store. Make a solution by dissolving five teaspoons of noniodized salt in one quart of water. Make another solution by dissolving ten teaspoons of noniodized salt in one quart of water. Pour the two solutions into separate shallow baking pans. Label each pan so that you will know which solution it contains. Put the two pans where they will remain at room temperature and will not be disturbed. Sprinkle a small amount of dried shrimp eggs (about 1/12 of a teaspoon) on the water in each pan. Every couple of days, add a bit of lettuce and a pinch of yeast suspended in fresh water to keep your shrimp alive. Every day for seven days, count the number of live hatched shrimp in each of ten teaspoons of water from each pan. (You will probably need to use a hand lens to see the shrimp clearly.) Find the average count per teaspoon. Now find the estimated total population in each pan by multiplying the average for each pan by 172 (the number of teaspoons in a quart). Record this information daily, as shown in the diagram.

	Days 1	2	3	4	5	6	7
Tsp. 1							
2							
3							
4							
5							
6							
7							
8							
9							
10							
Total							
Average per tsp.							
Average per pan							

At the end of the seventh day, plot your population information on a graph. Indicate the number of live hatched shrimp on the vertical axis and the number of days on the horizontal axis. Plot your data with two colored pencils, one color for each population. Compare the two curves. What kinds of curves are



they? How are they alike? How are they different? Why? Did the population in either pan decline? If so, on which day did it begin to decline? During the seven days what differences, if any, were there in the appearance of the shrimp in the two pans?

Now Think About This

Using what you have read and done, think about the problem below. Then write your answer, and explain it, under Part V of your record page.

In the classroom experiment, if you had used fresh water instead of a salt solution,

A more brine shrimp would have hatched from the original eggs

B the brine shrimp would have hatched more rapidly

C there would have been fewer live brine shrimp at all times

APPENDIX D

The three tables which follow present the regression equations used in the major analyses. In each table the full models are reported with the calculated constant, and regression weights for the respective vectors.

The equations are reported for the criterion measure X_3 (post STEP test), using as the restricted model, model 2.

If model 4 is used as the restricted model in the regression analysis, then the criterion measure predicted would be X_4 (teacher constructed post-test).

TABLE XXIV

REGRESSION MODELS FOR TWO CRITERION MEASURES

FULL GROUP ANALYSIS

Hypothesis	Full Model	
1	7	$X_3 = 61.98 + 0.34\ X_2 + 0.36\ X_5 + 0.32\ X_7 + 0.36\ X_8 + 2.54\ X_9 + 0.0\ X_{10}$
1a	5	$X_3 = 99.98 + 0.26\ X_2 + 0.29\ X_5 + 0.26\ X_7 + 0.41\ X_8 - 5.09\ X_{14} - 5.95\ X_{16}$
1b	1	$X_3 = 84.23 + 0.30\ X_2 + 0.28\ X_5 + 0.25\ X_7 + 0.43\ X_8 + 7.78\ X_{13} + 3.65\ X_{15}$
2	9	$X_3 = 68.66 + 0.33\ X_2 + 0.33\ X_5 + 0.29\ X_7 + 0.40\ X_8 + 6.27\ X_{13} + 0.70\ X_{14}$
3	11	$X_3 = 79.88 + 0.30\ X_2 + 0.31\ X_5 + 0.29\ X_7 + 0.42\ X_8 - 0.62\ X_{15} - 4.47\ X_{16}$
Restricted models		
	2	$X_3 = 69.74 + 0.30\ X_2 + 0.39\ X_5 + 0.33\ X_7 + 0.31\ X_8$
	4	$X_4 = -19.68 + 0.12\ X_2 + 0.09\ X_5 - 0.02\ X_7 + 0.04\ X_8$

Vectors are listed on page 40

Hypotheses appear on page 41

TABLE XXV

REGRESSION MODELS FOR TWO CRITERION MEASURES

TEACHER ONE

Hypothesis	Full Model	
1	7	$X_3 = 128.11 + 0.05\ X_2 + 0.42\ X_5 + 0.38\ X_7 + 0.33\ X_8 + 3.89\ X_9 + 0.0\ X_{10}$
1a	5	$X_3 = 160.01 + 0.01\ X_2 + 0.34\ X_5 + 0.29\ X_7 + 0.37\ X_8 - 2.47\ X_{14} - 4.64\ X_{16}$
1b	1	$X_3 = 148.98 + 0.00\ X_2 + 0.39\ X_5 + 0.32\ X_7 + 0.37\ X_8 + 6.29\ X_{13} + 0.18\ X_{15}$
2	9	$X_3 = 139.7 + 0.02\ X_2 + 0.40\ X_5 + 0.34\ X_7 + 0.38\ X_8 + 7.12\ X_{13} + 1.67\ X_{14}$
3	11	$X_3 = 133.39 + 0.51\ X_2 + 0.39\ X_5 + 00.37\ X_7 + 0.34\ X_8 - 3.03\ X_{15} - 4.49\ X_{16}$
Restricted models		
	2	$X_3 = 150.67 + 0.00\ X_2 + 0.42\ X_5 + 0.35\ X_7 + 0.29\ X_8$
	4	$X_4 = -15.47 + 0.06\ X_2 + 0.10\ X_5 + 0.04\ X_7 + 0.07\ X_8$

Vectors are listed on page 40

Hypotheses appear on page 41

TABLE XXVI

REGRESSION MODELS FOR TWO CRITERION MEASURES

TEACHER TWO

Hypothesis	Full Model	
1	7	$X_3 = -56.73 + 0.59\ X_2 + 0.42\ X_5 + 0.53\ X_7 + 0.55\ X_8 + 0.00\ X_9 + 0.00\ X_{10}$
1a	5	$X_3 = -20.40 + 0.53\ X_2 + 0.36\ X_5 + 0.47\ X_7 + 0.52\ X_8 - 11.29\ X_{14} - 4.47\ X_{16}$
1b	1	$X_3 = -22.96 + 0.57\ X_2 + 0.22\ X_5 + 0.39\ X_7 + 0.67\ X_8 + 9.89\ X_{13} + 5.44\ X_{15}$
2	9	$X_3 = -59.10 + 0.68\ X_2 + 0.33\ X_5 + 0.48\ X_7 + 0.51\ X_8 + 5.48\ X_{13} - 7.97\ X_{14}$
3	11	$X_3 = 33.13 + 0.52\ X_2 + 0.38\ X_5 + 0.49\ X_7 + 0.61\ X_8 + 1.23\ X_{15} - 1.66\ X_{16}$
Restricted models		
	2	$X_3 = -56.73 + 0.59\ X_2 + 0.42\ X_5 + 0.53\ X_7 + 0.55\ X_8$
		$X_4 = -18.94 + 0.17\ X_2 + 0.05\ X_5 - 0.09\ X_7 - 0.03\ X_8$

Vectors are listed on page 40 Hypotheses appear on page 41